

A Conceptual Model for a Programme Monitoring and Evaluation Information System

by

Gabriel Komakech

*Thesis presented in partial fulfilment of the requirements for the degree
Masters of Philosophy in Social Science Methods at the
University of Stellenbosch*



Supervisor: Prof. Johann Mouton
Faculty of Arts
Department of Sociology and Social Anthropology

December 2013

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

December 2013

Copyright © 2013 University of Stellenbosch
All rights reserved

ABSTRACT

Literature on monitoring and evaluation acknowledges the complexity in the field. Many evaluation studies require empirical evidence to be integrated with decisions on standards and values to reach robust evaluative conclusions. In this context, organizations face a number of difficulties in attempting to develop computerized software for monitoring and evaluating their programmes. The situation is exacerbated by the lack of literature on how various concepts used in programme monitoring and evaluation could be arranged into a coherent pattern of concepts upon which the development of monitoring and evaluation software could be contingent. The aim of this thesis is to present a conceptual model for a programme monitoring and evaluation information system that can guide programme agencies in the procurement, design and development of software for programme monitoring and evaluation. The conceptual model is based on an assessment of several key concepts that characterize programme monitoring and evaluation: programme goals and objectives; programme activities; programme providers; administrators; funders; community stakeholders; macro-environment and relationship between them; personal goals and objectives; existing conditions; targeted individual (s); family friends, and community; macro-environment and relationships between them; programme participation and programme outcomes. Using purposive techniques, 15 relevant monitoring and evaluation documents were selected from within 3 large-scale programmes implemented in Uganda. These documents were used to identify and describe the features and attributes associated with each of the key M&E concepts.

The findings reveal that only eleven of the key concepts listed above were used by the three case study programmes. In particular, their use was geared mainly towards the collection of empirical evidence to demonstrate programme accountability requirements. The study arranged the eleven distinctions into a framework comprising of three dimensions: (1) programme design; (2) programme implementation plan; and (3) programme implementation result. The programme design dimension comprises of five key concepts used to capture the essential information on programme design. The implementation plan dimension comprises of three key concepts used to capture the essential information on the actions that have been planned by each programme. The implementation result comprises of four key concepts that capture the essential information on the outcome of both routine and terminal monitoring and evaluation activities.

ACKNOWLEDGEMENTS

To start out, I would like to admit that everything that happened; both preceding and resulting into the intellectual output presented in this thesis has been the work of the Almighty God. For this, I am grateful because it was his graciousness that led me to a workshop organized by Evaluation Research Agency (ERA) in late 2005. Two defining events occurred at the workshop:

(1) I learnt of an inaugural post-graduate diploma programme in monitoring and evaluation methods offered by Stellenbosch University.

(2) I met Professor Johann Mouton.

Meeting Prof. Johann Mouton was indeed a defining moment for me and one for which am forever grateful and deeply indebted. First, he gave me the opportunity to pursue a career of my long-time dream; monitoring and evaluation - beginning with the postgraduate diploma right up to this candidature. Secondly, he has been my supervisor and the greatest contributor without whom this work would not be. For this am very much grateful and would like to say a big THANK YOU.

I would also like to greatly thank my colleague, classmate, friend and business partner, Apolo Kyeyune with whom I have shared numerous intellectual arguments. Your critical and sometimes provocative critique helped to refine my thinking; led me to consult additional literature and to re-engage in our never-ending discussions on evaluations.

I am very thankful to the Carnegie Corporation of New York who offered the scholarship for both the postgraduate diploma and the Masters programme. Your support has kept me focused on my studies.

I would like to also extend my sincere thanks to the following staff of CREST without whom my numerous travels and stay at Stellenbosch would not have been as wonderful: Ms. Marthie Van Niekerk, Ms. Lauren Wildschut, and Ms. Charline Mouton. I sincerely appreciate all your contributions.

And finally, I am very fortunate to have a wonderful wife and children who all stood by me and supported me during the entire period of my study. I am particularly indebted to my wife, Lucy Alal Komakech who on many occasions was both 'the husband and the wife'. To my children Nicholas and Karen – who found a way of relieving my stress particularly during the course of writing the thesis with such statements as "daddy, your teacher has given you all this home-work!" The twins: Jordan and Jayden, who disapproved scientific instrumentations to be born twins and continue to be a blessing to the family and for giving me an additional name: "Salongo".

TABLE OF CONTENTS

ABSTRACT	i
ACKNOWLEDGEMENTS	ii
1 CHAPTER 1: INTRODUCTION	1
1.1 Background.....	1
1.2 Monitoring and Evaluation software landscape	4
1.2.1 Software for data identification	4
1.2.2 Data Collection Software	4
1.2.3 Data Analysis Software	7
1.2.4 Reporting Software	8
1.2.5 Adaptable Software	8
1.3 Description of selected software	8
1.3.1 PEPPMIS.....	9
1.3.2 PEMS.....	10
1.3.3 LOGICS	11
1.3.4 eM&E™	13
1.3.5 CRIS	15
1.4 Research problem and objective	16
1.4.1 Statement of the problem	16
1.4.2 Research objective	17
1.5 Design and Methodology	18
1.6 Layout and structure of the thesis	18
2 Chapter 2: LITERATURE	20
2.1 Introduction	20
2.2 Classification of evaluation models	20

2.3	Evert Vedung	21
2.3.1	Effectiveness model	22
2.3.2	Economic model	30
2.4	Daniel Stufflebeam	32
2.4.1	Question/Method-Oriented Studies	33
2.4.2	Improvement/Accountability-Oriented Studies	43
2.4.3	Social Agenda/Advocacy Approaches	47
2.5	Concluding comments	52
3	CHAPTER 3: ANALYTICAL FRAMEWORK	54
3.1	What is a system?	54
3.2	Systems thinking and programme evaluation.....	56
3.3	DSRP Patterns for a human service programme	59
3.3.1	The provider system	59
3.3.2	The target system.....	64
3.3.3	The human service system.....	65
3.4	Conceptual modeling theory	66
3.5	The Analytical Framework	66
4	CHAPTER 4: DESIGNN AND METHODOLOGY	69
4.1	Unit of Analysis	69
4.2	The Analytical framework	70
4.2.1	Distinction	70
4.2.2	Attribute	70
4.2.3	Relation.....	71
4.3	Sampling	71
4.4	Data collection and analysis	73
4.5	Limitation of the study.....	74

5	CHAPTER 5: RESULTS.....	75
5.1	The AIDS/HIV Integrated Model District Programme	75
5.1.1	Description of distinction.....	76
5.1.2	Description of Relation	82
5.2	The Northern Uganda Malaria AIDS Tuberculosis Programme (NUMAT)	84
5.2.1	Description of distinction.....	84
5.2.2	Description of Relation	89
5.3	The President's Emergency Plan for AIDS Relief (PEPFAR)	90
5.3.1	Description of distinction.....	91
5.3.2	Description of Relation	100
5.4	Concluding comments	101
6	Chapter 6: Discussion	103
6.1	Programme goal	103
6.2	Programme Objective	106
6.3	Programme Activity.....	108
6.4	Programme Provider.....	110
6.5	The PMP and the COP	111
6.6	The Reflector Conceptual Model	112
7	Chapter 7: Conclusions and recommendation	115
8	REFERENCES.....	118

List of Tables

Table 1: Evaluation descriptors	21
Table 2: Characteristic of effectiveness approaches on organizer & stakeholder descriptors	29
Table 3: Characteristic of question/method approaches on organizer & stakeholder descriptors	42

Table 4: Characteristic of question/method approaches on organizer & stakeholder descriptors	47
Table 5: DSRP pattern for a human service programme.....	61
Table 6: Case study programme and the documents sampled	72
Table 7: PEPFAR legislative goals and indicator (source: Next generation indicator reference guide).....	91
Table 8: Country summary for PMTCT legislative goals and indicator	93
Table 9: Summary of distinctions across AIM, NUMAT and PEPFAR	101
Table 10: PEPFAR legislative goals and indicator.....	104

List of Figures

Figure 6.1: Conceptual Model for Programme goal	105
Figure 6.2: Conceptual model for programme objective.....	107
Figure 6.3: Conceptual Model for Programme Activity	109
Figure 6.4: Conceptualizing Programme Provider.....	110
Figure 6.5: Conceptual Model for Programme Work plan	111
Figure 6.6: The Reflector Conceptual Model	113

CHAPTER 1: INTRODUCTION

This chapter contextualizes the research question and provides justification for undertaking the study. It explores various ways that software is used in monitoring and evaluation; and identifies the problem that the study is addressing, along with justification. An appropriate design for researching into the problem is presented, and finally, a description of the layout and structure of the thesis is provided.

1.1 Background

A major event that influenced the choice of question for this thesis was an initiative in 2002 to develop monitoring and evaluation software for the Uganda's ministry of local government (MoLG). The initiative resulted in a software product known by the acronym LoGICS¹. Towards the end of the assignment, a viewpoint emerged that the design of LoGICS could constitute a framework for developing generic software for monitoring & evaluation (M&E).

This viewpoint was put to test in 2006 when I was leading a team of software developers tasked with re-designing and clearing LoGICS of bugs². I used the opportunity and attempted to redesign LoGICS with enough flexibility. The intention was to evolve it into generic M&E software but although the resulting product was a considerable improvement over the previous version, its adaptability to different settings other than MoLG was not achieved. Even within MoLG, it still had limitations and could not be extended to cover every scenario. At the end of the assignment, just like it was at the beginning, one question remained un-resolved:

How can software for monitoring and evaluation be designed to allow adaptability across different programmes?

Although not apparent at the time, this earlier failure to develop generic software for M&E was a result of the complexity in the M&E field itself. Monitoring and evaluation is a field known to embody many intricacies: Bulgarelli and Gori (2004) argue that the ability of outsiders to understand the concepts of M&E is limited by the multiplicity of definitions, formalizations and measures used by evaluation specialists. This viewpoint is shared by Crawford who identified three conceptual issues hindering the practice of M&E:

(a) ambiguity in the definition of monitoring and evaluation;

¹ Local Government Information and Communication System

² Errors in software program

(b) divergent philosophical views about how the change anticipated by aid projects may be both represented and judged;

(c) the various perspective from which a monitoring and evaluation information system (MEIS) may view the performance of 'the project' vis-à-vis the performance of 'the implementing agency'. (2004:142)

In a quest to satisfy a long-time interest in M&E software I began to explore avenues that would better my understanding in M&E. There is recognition that the success of software projects hinges partly on shared knowledge held by software developers and application domain experts. In a recent study; Tesch, Sobol, Klein and Jiang (2009) concluded that a combination of both user knowledge of information system (IS) development and IS developer knowledge of application domains had significant impact on successful project outcomes. Consequently, projects where the developers possess application domain knowledge are likely to be more successful. Fortunately for me, I learnt of an inaugural postgraduate diploma programme in Monitoring and Evaluation Methods (MEM) at the University of Stellenbosch. I enrolled and subsequently graduated in early 2007. Afterwards, I continued into this candidature, through enrolment into a Master of Philosophy programme in Social Science Methods (MPhil SSM). This was motivated by two issues: first, given my background in the sciences³, I was challenged with having to apply methodologies of the social sciences, a common practice in M&E (Rossi, Lipsey & Freeman, 2004, pp.16). Secondly, the candidature offered opportunity to deeply investigate the issues concerning development of adaptable M&E software.

Prior to this candidature, and during its early phase; I considered a major output of the research to be adaptable M&E software. Although still personate about the idea, the ultimate focus of the research shifted from adaptable M&E software being the envisaged output to a conceptual model for the development and application of such software.

There are several approaches for developing conceptual models however in this thesis, the approach chosen is the conceptual modelling approach. According to Juristo and Moreno (2000), conceptual modelling⁴ has gained importance in situations where the problem to be solved is located in a domain that is further away from the software developer. In other words, the software developer possesses very little knowledge in the problem domain. In keeping with this notion, this thesis aimed to *represent* the information requirements of a monitoring and evaluation information system (MEIS) in a way that enhances software developer's ability to *understand* and thus be able to build adaptable M&E software system. Since such a software system captures, stores and processes information of some real world situation, a valid representation of the real-world is needed if the software is to be useful to its end-users. Therefore, the conceptual model presented in this thesis is a representation of the practice and theory of M&E.

³ The author, in addition, holds Bachelor of Science and Master of Science degrees in computer science.

⁴ The process of creating Conceptual Model in software development is generally referred as conceptual modeling

There are a number of practical reasons underpinning the decision to develop a conceptual model for MEIS;

- i. According to Niehaves, Ribbert, Dreiling and Holten (2004), a major reason for the failure of IT projects is a miscommunication between business and IT personnel. In the views of Niehaves *et al*, this is a result of a paradox:

Business personnel are not usually able to explicitly give their information requirements to IT personnel in a way that can be technically used to implement or configure a system. IT personnel, on the other hand, usually, do not have a business background detailed enough to provide business personnel with appropriate IT solutions independently. (2004:4232)

Thus, a conceptual model for MEIS is viewed as a means to: support the communication between software Developers and Customers/or Users; help software developers and analysts understand the MEIS domain; provide input for the design of MEIS software; aid documentation of the original requirements to be used for future reference (Dieste *et al*, 2004:5);

- ii. Developing off-shelf-applications: a conceptual model provides a general description of the structure and behaviour of a MEIS. In this way, a conceptual model for MEIS is an attractive artefact for developing off-shelf software for M&E (Fettke & Loos, 2007);
- iii. Communicating best-practices: constructing a conceptual model involves interaction with domain experts. In the process, best practices can be embedded in the resulting model, thus fostering the development of high-quality software (Fettke & Loos, 2007);
- iv. Selection/specification of M&E software: a conceptual model is an important artefact for any organisation that may want to procure or develop software for their M&E operations. In such situations, the model is a starting point in specifying the software requirements. The benefits associated with this are cost and time saving (Fettke & Loos, 2007).

As previously mentioned, the study began with the ambition of developing adaptable M&E software, but ended with a conceptual model for MEIS. The unfolding and refocusing was a result of literature review, paying particular attention to review of literature about the different kinds of software being employed in the conduct of M&E. This is a central issue of discussion in the next section.

1.2 Monitoring and Evaluation software landscape

There are various types of software found to support monitoring and evaluation activities. These software programs either support specified phases of the M&E data cycle or a combination of phases. The M&E data cycle comprises of six phases namely (Crawford, 2004): data identification, data capture, data analysis, dissemination, utilization and assessment.

1.2.1 Software for data identification

Evaluation planning is an initial, albeit important first step in many evaluation studies. It is concerned with identification of the criteria on which value judgements are based and the corresponding data elements required to answer the evaluation questions. This thesis posits that the software for data identification can be grouped roughly into two categories: (1) general purpose software and (2) specialized planning software.

Many of today's computers ship with software such as word processors, spreadsheets, presentations, browsers, electronic mail (e-mail), and groupware already pre-installed. These kinds of software are frequently used to facilitate evaluation planning and dissemination. Such use has been demonstrated by Leslie, Holosko and Dunlop (2006) to include: searching the Internet for previously tested / validated instruments and relevant literature; using e-mail to disseminate planning meetings, ongoing reports/minutes and facilitate feedback and follow-up discussions. In contrast though, the specialized planning software is intended for use in supporting planning activities such as strategic planning, M&E planning and performance management planning. An example of specialized planning software discussed in this thesis is DoView (<http://www.doview.com>).

DoView allows evaluators to quickly produce visual models of the outcomes that a programme or project is trying to achieve and the steps involved to achieve those outcomes. Such visual models go by many names such as: outcomes models, results models, strategy maps, logic models, intervention logics, theories of change, programme theories and ends-means diagrams (DoView, 2009). DoView is particularly useful in identifying data for an evaluation because its visual models are made up of the essential elements required in organizing an evaluation study. Elements in a DoView model include question, input, activity, output, outcome, objective, goal, indicator and service element. The DoView website provides several examples of monitoring and evaluation plans developed using DoView.

1.2.2 Data Collection Software

There are various ways of leveraging information technology to collect evaluation data. In such instances, an electronic form is loaded onto a computer or handheld device. Respondent and field

enumerator use the electronic form to electronically capture the required data. This technique of data collection includes web-based surveys; mobile data collection and e-mail-based data collection.

The first variant of data collection software provides web-based electronic forms. Respondents are invited to visit the website and to complete the survey. This approach was found to be the most widely cited means that software is used in M&E data collection. There are several software products found to support web-based surveys (Crawford, 2002; Wright, 2005)⁵: InstantSurvey (<http://www.instantsurvey.com/>); SPSSMR (<http://www.spssmr.com/>); SurveyMonkey (<http://www.surveymonkey.com/>); Survey Solution (<http://www.vovici.com/>); Zoomerang (<http://www.zoomerang.com/>). The major functionalities of this type of data collection software are (Crawford, 2002; InstantSurvey, 2009; SPSS Inc., 2009; SurveyMonkey.Com, 2009; Vovici Corporation, 2009; Zoomerang, 2009):-

- i. *Survey creation tool* – This provides tools for designing the survey questions, and may include a questionnaire wizard that guides the user through an automated process of creating a survey. The creation tool supports various question types - single choice, multiple choice, matrix, numeric entry, text, memo, constant sum, pull-down and custom question styles; and basic text editing – such as being able to manipulate fonts, colours and pictures; and also logic check such as behind-the-scene variable calculations, filling text responses into later parts of the survey, dynamic creation of response options based on previously provided response; validation capabilities such as mandatory responses, comparing responses against other responses or preloaded data, range checks, data format checks;
- ii. *Survey templates* – This provides a collection of already designed and tested surveys (questionnaires), which the evaluator can adopt or use as a basis for constructing his own survey;
- iii. *Respondent management* – This offers tools for managing respondents, with common management functions such as adding new respondents (including loading of their e-mail contact addresses), inviting respondents to complete a survey and removing respondents from the system;
- iv. *Report* – This provides reporting tools that may be used to query the survey data and generate simple statistics such as frequencies, graphs/charts, measure of central tendency, cross-tabulations and to export the survey data into formats compatible with most data analysis software. This reporting function is only appropriate for preliminary viewing of the data and generation of simple statistics. For more detailed and advanced analysis, the specialized data analysis software is still required.

⁵ There are many more web-based survey tools available in the market than are listed here. It should be noted that the listing in this thesis is not based on any ranking, and should therefore be used as is.

The second variant of the data collection software provides the possibility of collecting data onto a form which has been loaded onto a hand-held device, most often a personal digital assistant (PDA) or a mobile phone. There are several examples of mobile data collection software that the thesis identified: Open X Data (<http://www.openxdata.org>); Frontline SMS/Frontline Forms (<http://www.frontlinesms.com>); Mobile Researcher (<http://www.populi.net/mobileresearcher/>); EpiSurveyor (<http://www.datadyne.org/epSurveyor/>); Nokia Data Gathering (<http://www.nokia.com/corporate-responsibility/society/nokia-data-gathering/english/>); Open Data Kit (<http://www.opendatakit.org>); Emit (<http://www.emitmobile.co.za>); EpiCollect (<http://www.epicollect.net>); Voxiva (<http://www.voxiva.com>). This variant of data collection software allows only the reporting staff or enumerator to access the electronic form, unlike in the web-based variant where respondents also have access to the online form. This is very much the case of an interviewer-administered survey design, except that the questionnaire is loaded onto a PDA or mobile phone. The major functionalities of this variant of data collection software can be grouped as follows:-

- i. *Survey creation*– Designing and loading the electronic form onto a handheld device is performed by technical personnel, unlike in the web-based variant where the web form can be designed by the evaluator. Form features such as question types, basic text formats, logic checks and validations are hard-coded into the form using programming tools and languages;
- ii. *Synchronization* – This feature facilitates the transfer of survey responses from the handheld device to a computer. Synchronization happens whenever the handheld device is connected to the computer. Connectivity between the PDA and the computer is either through a local or remote connection and is dependent on the distance between the PDA and the computer. A local connection is achieved by attaching the PDA directly onto the computer using a special cable while remote connectivity is achieved by attaching the PDA to a mobile telephone provider's network;
- iii. *Report* – The PDA does not usually provide ability to generate report. However, reports can be generated from the destination computer where data from the PDA is sent.

A third variant of data collection software provides for the possibility of collecting data via electronic mail (email). In this technique, an electronic form is created with all the necessary questions that respondents must complete. The form is then sent via email to respondents and upon receipt, it is displayed for them to enter their responses. When the electronic form is filled and sent back; the contents of the form are automatically added to the appropriate data repository – thus eliminating manual data entry. Two examples of email-based data collection software identified in the study are Eform (<http://www.beachtech.com>) and a combination of Microsoft Access 2007 and Microsoft Outlook 2007 (<http://www.microsoft.com>).

1.2.3 Data Analysis Software

The data analysis software is concerned with the process through which data that has been captured is subjected to some form of treatment, transformation or contextualization in order to derive meaning. There are two categories of data analysis software; software that facilitates analysis of quantitative data and software that facilitates analysis of qualitative data.

Software that is designed to facilitate quantitative data analysis possesses features for the execution of advanced descriptive and inferential statistical operations. Such software allows data files to be loaded from a variety of file types – such as relational database files, spreadsheet files and text delimited files, and can operate in both a standalone mode (the software is installed on a user's computer) or network-based mode (the software is installed on a central computer on the network from where other users connect and use the system). The predominant products under this category are: SPSS, <http://spss.com>; EPI Info, <http://epi-info.com>; SAS, <http://sas.com>; STATA, <http://www.stata.com>; STATISTICA, <http://www.statsoft.com>. These products are fairly mature and have numerous books and manuscripts dedicated to them.

Software that is designed for qualitative data analysis provides functionalities for managing texts⁶ and their coding; examining how frequently and how words are used in context as well as exploring the coding, e.g. how often particular categories have been assigned to a word or text segment, which categories and how often they occur, what links or relations exist between categories or coded text segments; creating and maintaining categories and categorisation schemes; assigning one or more categories/codes to word strings, words, phrases, sentences, lines, paragraphs or whole texts; keeping notes (memos) on text, categories, coded text segments; obtaining different views of the text data as well as the coded parts of a text or a group of texts; exporting the coding for further processing with other software as well as generating reports on the performed analysis and supporting team or co-operative work for a text analysis project and merging codes (Alexa & Zuell, 1999).

There are many software programs designed to assist in qualitative data management. The most commonly cited ones include (Alexa & Zuell, 1999): Atlas.Ti, <http://atlasti.com>; Ethnograph, <http://www.qualisresearch.com>; HyperResearch, <http://www.researchware.com>; MAXqda, <http://www.maxqda.com>; NU*DIST, <http://www.qsr.com.au>; NVIVO, <http://www.qsr.com.au>.

⁶Depending on the software, managing and coding multimedia and audio material or support for their transcription process is also possible

1.2.4 Reporting Software

The reporting software category is concerned with the production/dissemination of reports required by different stakeholders. In its typical form, reporting software enables programme/project implementers to capture service data for their own use as well as for summary and analysis by other stakeholders, such as funders. The intent is to standardize data collection and reporting for an entire programme or policy.

The reporting software is mainly found in situations where a “master-subject” relationship exist. The “master” (“principal”) is an organisation that supports/funds multiple grantees to implement its programme interventions to a target audience who are usually located at various sites. At each site, a grantee (termed “subject”) is responsible for delivering the programme interventions. The “master” organisation provides standardized guidelines to govern issues such as the interventions/services and its delivery; monitoring and reporting schedules; data collection templates and their schedules; periodic evaluation studies and schedules. The reporting software is designed to automate the guidelines – and is therefore developed in accordance with guidelines in both look and functionalities. While several examples of software products under this category may exist, a discussion of three such software is provided: The Local Government Information and Communication System (LoGICS); The Performance Evaluation and Monitoring System (PEMS) and the President’s Emergency Plan Performance Management Information System (PEPPMIS). Detailed discussions of PEPPMIS, PEMS and LoGICS are provided in sections 1.3.1; 1.3.2 and 1.3.3 respectively.

1.2.5 Adaptable Software

Adaptability is a non-functional requirement of software products along with security, performance, maintainability, reusability, support, training and documentation. It deals with the extent to which a software system adapts to changes in its environment. Adaptable software has features that allow its behaviours to be adjusted without the need for re-programming (Stiemerling, Kahler & Wulf, 1997).

Two examples of adaptable M&E software are eM&E™ and the Country Response Information System (CRIS). Software under the adaptable category automates many features of the M&E data cycle (see section 1.2). Functionalities of such software include support for data identification, data collection, data analysis, reporting and dissemination. For a detailed discussion of eM&E™, and CRIS; see sections 1.3.4 and 1.3.5 respectively.

1.3 Description of selected software

In the previous section, five categories of software that are commonly used in M&E are highlighted. From the discussion, it is evident that software belonging to the first three categories (general

purpose, data collection, data analysis) is not necessarily M&E software. For instance, software in the data collection category may be used by non M&E studies that require to collect some data in which case, anybody interested in electronic data collection can make use of it, irrespective of whether or not the data being collected is for M&E purposes. The same argument holds for both the general purpose and the analysis software categories. It was noted that software under the reporting and adaptable categories are designed exclusively for supporting M&E activities. In this section, five software products under the reporting and adaptable categories are described in some detail.

1.3.1 PEPPMIS

The President's Emergency Plan Performance Management Information System (PEPPMIS) is software developed to support the monitoring and evaluation efforts of the President's Emergency Plan for AIDS Relief (PEPFAR) programme. Its main objective was to strengthen the collection, storage, merging, sharing and reporting of PEPFAR data among United States Government (USG) agencies⁷ and Seventy-Six (76) USG-funded partners (Moon & Smith, n.d). The software, then in its seventh release, was first used in 2005, but had to be enhanced to cater for the PEPFAR reporting requirements of 2007: The President's Emergency Plan for AIDS Relief Indicators, Reporting Requirements and Guidelines, July 2007 (MEEP Project, 2008).

The description of PEPPMIS is based on various documents sourced from the Internet including: MEEPP Project, 2008; Moon & Smith, n.d; PEPFAR, 2008. PEPPMIS is developed using Microsoft technologies and is accessed using web-browsers. The recommended browser is Internet Explorer version 7 or higher. The system has security features that require users to provide login name and password. Partners are only allowed to enter data during a "window" of data entry and thereafter, all users can only view data. When the data entry "window" is closed, Implementing Partners are notified in writing regarding data anomalies. Any clarification requiring data update has to be accompanied with an authorization order that grants permission to the central authority to update the Partner's data. Once all issues dealing with data entry are cleared, data from different Implementing Partners is aggregated centrally for further processing and reporting.

The core functionality of PEPPMIS is organized around the themes of prevention, care, treatment and workforce. Within the four themes, a minimum set of 46-programme-level indicators are prescribed, and each partner is obliged to collect and report data on each. PEPPMIS is designed to capture data for the 46-programme-level indicators, and to generate corresponding reports. For each indicator, there are four different types of data that is collected: The number of organizations provided with technical assistance; the number of service outlets assisted; the number of clients served and; the

⁷ Department of State, United States Agency for International Development (USAID), Centre for Disease Control (CDC), National Institute of Health (NIH), Department of Defence (DoD) and Peace Corp

number of people trained. Besides the indicators, PEPPMIS does allow partners to report on funds obligated. A major benefit reported of the system is the enforcement of compliance and standardization in data collection. The software made it possible to consolidate and aggregate data from all implementing partners, thus presenting a national perspective of the HIV/AIDS epidemic. System support was provided centrally through the office of the monitoring and evaluation of the emergency plan progress (MEEPP) located in Kampala.

1.3.2 PEMS

To facilitate the monitoring of HIV prevention programmes, the Center for Disease Control (CDC) in Atlanta, USA developed the Programme Evaluation and Monitoring System (PEMS). PEMS was a national data reporting system that comprised of a standardized set of HIV prevention data variables, secure web-based software for data entry and management, and a range of data collection training and software implementation support services (Thomas, Smith, & Wright-De Agüero, 2006). PEMS data enables HIV prevention stakeholders at all levels to examine programme fidelity and to monitor key programme health service utilization and behavioral outcomes. In addition, PEMS enables CDC to identify best practices and assist grantees in redesigning interventions to accomplish stated goals such as the reduction of high-risk behaviors in targeted populations. Finally, the PEMS data can be used to compliment other data collection systems such as behavioral surveillance, HIV/AIDS surveillance, and special studies projects to better monitor prevention efforts and the epidemic from the local and national perspective. Typical questions that can be answered from PEMS include (Thomas, 2008):

- i. Programme-monitoring questions: *what are the characteristics of HIV prevention programme as planned; what are the characteristics of prevention programme as delivered; to whom were HIV prevention programme provided? What resources are allocated to these services? What behavioural and service utilization outcomes do client reports?*
- ii. Programme-evaluation questions: *to what extent is the programme reaching its intended target population? To what extent is the programme plan being delivered as intended? To what degree are the programme performance indicator targets being achieved?*

Collection and reporting of the PEMS data set was a requirement for all health departments and community based organizations (CBOs) funded through CDC HIV prevention cooperative agreements. The PEMS web-based reporting software is implemented based on a standardized set of data variables (CDC, 2008).

The data is collected around five themes (Thomas, 2005):

- i. Agency Characteristics—budget, sites, workers, contracts, & network agencies;
- ii. Programme Plans – programme models, target populations, interventions, settings, sessions, & activities;
- iii. Client Information—demographics, risk profile, detailed risk behavior assessments;
- iv. Service Delivery—service activities, recruitment, and referrals;
- v. Community Planning—Target populations and priority interventions.

There are a number of concerns that have been voiced concerning PEMS. These are generally non-technical and include (CHAMP, 2005):- PEMS data collection is extensive and unduly burdensome. The PEMS data set is 228 pages long, and prevention staff would be required to conduct an interview with each and every client during each and every CDC-funded encounter. As CHAMP quoted one disgruntled staff:

“PEMS is not going to evaluate our intervention. It is going to be our intervention” (2005:2).

1.3.3 LOGICS

The Local Government Information and Communication System (LOGICS) is software developed by Uganda’s Ministry of Local Government (MoLG) to support evidence-based planning and decision making in Local Governments (LGs). The aim was to empower LGs with tools that could allow them to better monitor & evaluate the level and quality of service delivery; the progress of district-level projects; and to assess/enforce operational compliance with established laws and regulations. LoGICS was first introduced at LGs in 2002, and following recommendation from numerous studies (Chalmers, 2005), the entire software was overhauled in 2006 to introduce new functionalities and address short-comings.

The LoGICS software is developed using Microsoft .NET technology and includes separate data entry and a reporting components that can run on any personal computer (PC) installed with Microsoft Windows operating system. The reporting component is a web-based service, and can be accessed locally (within a local government) or from the Internet. The functionalities of LoGICS are provided in four distinct modules (SysCorp International, 2006):

- i. Service Delivery module: This module is used to track the extent to which LGs are performing. The system defined over 500 indicators to track performance across a broad range of sectors. The specific features of the service delivery module (the indicator system) are:
 - A flexible interface for adding and modifying indicators, which grants LGs the ability to collect and analyze emergent information requirements;

- Support for local indicators: In addition to the national-level indicators (mandatory indicators for all local governments), the system provides ability for local governments to define and track indicators that measure aspects local to them;
 - Support for calculated indicators: the system allows new indicators to be derived from existing ones based on mathematical formulae. Once the appropriate formula is defined, the indicator is automatically generated from existing data – eliminating the possibility of introducing error, which is characteristic of manual calculation.
 - Automatic generation of data collection instruments: Once all indicators are defined for a given service-delivery unit, the system can automatically generate the data collection form that incorporates all the indicators defined for that type of service-delivery unit (e.g. primary school or health facility). This can be printed and distributed to each of the service-delivery unit for filling, along with a corresponding instruction for completing the form, which is also generated from the system.
- ii. Project Cycle Management (PCM) module: The PCM module is used by LGs to track the implementation of projects defined in the annual work plans. The following categories of data are tracked:
- Project-plan-data: This functionality is used to capture data relating to the project plan. Examples: project summary (implementer; funder(s); location; approval log history); project cost; targeted beneficiaries;
 - Project-procurement-data: details relating to project tendering such as date of tender and date of contract award;
 - Project-progress-data: details relating to project implementation progress such as current vis-à-vis planned expenditure; actual vis-à-vis planned outputs;
 - Project-completion-data: project cost vis-à-vis planned cost; project financing vis-à-vis planned financing; project expenditure vis-à-vis planned expenditure; project beneficiaries' vis-à-vis planned beneficiaries.
- iii. Compliance Inspection: The compliance inspection module assesses whether or not LGs are operating in accordance with established laws and regulations. The core of the system is a set of questions and corresponding scoring scheme that measures performance of LGs against prescribed standards. The system is developed with sufficient flexibility that allows new standards to be added, along with its scoring scheme.
- iv. Web-based reporting facility: Reports in LoGICS are viewed using a web-based interface and is provided in two categories: standard and custom. Standard reports are available in pre-formatted form, with the formats pre-determined by the primary recipient of the report. Custom reports are generated dynamically based on the needs and selection of individual users. Data can be aggregated to give a national, regional, district, sub-district or facility level perspectives.

- v. One Stop Resource Center (OSRC): LoGICS is designed to operate as a ‘One-Stop Resource Centre’ within the MoLG. Data files from LGs are sent to the MoLG to be uploaded into a consolidated database which is accessed from the Internet using a web-based reporting facility.

1.3.4 eM&E™

The eM&E™ software is developed and marketed by Aid-IT⁸ Solutions; an IT company based in Australia. This software is considered the world’s first fully configurable MEIS, and is currently in version 2. The description of eM&E™ presented below is based on information gleaned from the product’s website (Aid-IT Solutions, 2007).

The eM&E™ software is developed using Microsoft .Net technology and can run on any personal computer (PC) that is installed with Microsoft Windows 2000 or later. The software operates in a mix of centralized and decentralized model.

The installations of eM&E™ software at user sites rely on a central server located in Australia for its complete functioning. At user sites, the software is run directly from a USB stick, without requiring any software installation on a user’s PC.

At periodic intervals, a connection is established between the central server and the computer on which the USB is attached. The connection is used to transfer project data from the USB to the central server, and any configuration information, if available, may be pushed to the USB.

Although not explicit, there is every indication that eM&E™ is developed in conformity with the *AIDING AID* framework. There are at least two reasons why this is the case. First, the principal consultant at Aid-IT, Dr Paul Crawford is the author of the *AIDING AID* framework – a framework he developed as part of his doctoral thesis⁹. Secondly, the functionalities in eM&E™ are organized around the presupposition advanced in the *AIDING AID* framework. The functionalities of eM&E™ include:-

- i. **Data variables:** There are different kinds of data that eM&E™ captures, analyzes and generates reports on. These data types mostly correspond to specific monitoring and evaluation functions being undertaken. A first category of data relates to project outputs. EM&E™ arranges project outputs by types, for instance, *hand-dug wells* is one type of output. Each output had several attributes attached to it. An output such as hand-dug wells may be associated with attributes such as depth, name, date started, number of community

⁸ <http://www.aid-it.com.au>

⁹ AIDING AID: A monitoring and evaluation framework to enhance international aid effectiveness. A doctoral dissertation by Paul Crawford

supplied laborers, GPS coordinates, flow rate and hand pump installed. The attributes are useful in monitoring and reporting about the completeness of each output, and are defined at system setup. The software focuses on tracking planned vis-à-vis actual outputs. A second category of data relates to project activities. The software provides ability for project staff to plan and capture activities that produced a given output. Again, using *the hand-dug wells* as an example, related activities may include deliver additional concrete materials to a given site; organize orientation for selected village laborers. These activities are captured in the software and their progress tracked. A third category of data relates to project surveys. The software provides ability for the project to plan, conduct, capture and analyze survey data. The survey data is also the primary source of information that eM&E™ uses to deduce effectiveness (effect data) of a project. A fourth category of data relates to risk data. Risk data is captured at three levels: management risks, intervention risks and development risks. At each level, the STEEP¹⁰ mnemonic is used to guide the selection of risks that should be captured. A fifth category of data relates to project financials. In project financials, the software places special interest in capturing planned vs. actual expenditure. A sixth category of data relates to project narrative report. The narrative reporting feature allows project staff to describe aspects of the project along pre-defined narrative categories. Narrative categories may include human resource, risks and general issues or concerns. A seventh category of data relates to activity feedback. The feedback feature allows supervisors to comment on progress that their subordinates are registering on activities, outputs and effects.

ii. **Data entry modes:** eM&E™ provides three avenues through which data can be captured:

- Keyboard - data from paper forms is entered directly into the system on the USB flash drive;
- A PDA - data is entered directly into mini version of the software loaded onto a PDA, and is later uploaded into the USB flash drive;
- Scanning software - data is scanned into the system using form recognition technology, with no typing required, eliminating another opportunity for human error.

There is limited write-up about eM&E™ even on its website. On the product's website, it is indicated that the software evolved through 16 version updates and has been implemented or trialled by 5 partner organisations across 12 projects in 8 countries in Africa, Asia and Australia.

¹⁰ Social Technological Economic Environmental Political

1.3.5 CRIS

The Country Response Information System (CRIS) is developed by the Joint United Nations Programme on HIV/AIDS (UNAIDS) as a database-supported information system to facilitate the collection, storage, retrieval and dissemination of a range of existing information. The core information in CRIS relates to HIV/AIDS indicators, resources, and scientific research (UNAIDS, 2003). In this way, CRIS helps to create a picture of the effectiveness of ongoing programmes and costs associated with a country's response to HIV and AIDS. A major assumption behind its development is an expectation that national governments will adopt it as a unifying platform to house all HIV/AIDS-related indicators that are being collected, irrespective of who collects the indicator.

The CRIS software is developed using Microsoft .NET technology and runs on Microsoft windows environment. End users are able to access the application using compatible web browsers, such as Internet explorer. The major functionalities in CRIS are delivered in three separate modules, namely (UNAIDS 2003; UNAIDS n.d):

- i. The indicator database: This module allows countries to collect and analyse indicators of the HIV/AIDS epidemic. Within the database, indicators are categorized into *core* and *free* indicators. The core indicators are pre-configured and ship with the system. They correspond to indicators that have been endorsed at international level. Modification to the core indicator list is done centrally by UNAIDS, with countries only being able to import it into their local installation. The *free* indicator facility allows countries to adapt CRIS to their local context. This facility grants countries the liberty to define indicators that measure unique aspect of the epidemic, and may vary from country to country. All data import/export is made possible through a data exchange facility that enforces consistency in data exchange formats. The functioning of the indicator database is further enhanced by the so-called Global Response Information Database (GRID), a web-based reporting portal. The GRID is an aggregated database containing indicators that are derived from the world-wide installations of CRIS. The GRID allows comparison or analysis of the HIV/AIDS epidemic between countries and within regions; and to provide a global picture of the epidemic. One aspect of the indicator database is that it is designed to be HIV-AIDS specific, and specifically to address the indicators that emerged from the UNGASS Declaration of Commitment on HIV/AIDS. In addition to these indicators, the later version of CRIS (Version 3) has added support for the PEPFAR indicators.
- ii. Project/resource tracking database: This database is complementary to the indicator database and is primarily intended to support improved national planning, resource mobilization/allocation, intervention targeting and evaluation and analysis of a country's success in implementing its own National Strategic Plan (NSP), and analysis of its efforts and

compliance with the UNGASS Declaration of Commitment on HIV/AIDS and other regional or global commitments. In other words, nations are able to analyse funding and programme gaps by any combination of time frame, geographic area, target population, type of project and organization. In addition, data in the system is coded by geographical locations to allow for analysis against other data, such as census data, school attendance data, health data, transportation data and agricultural data. Data in the system can be reported in a multiplicity of dimensions: sub-national level, such as province or district; executing/implementing organization or type of organization (government ministry, provincial ministry, UN agency, NGO); resource provider (donor); planned or actual start and/or end dates for projects; project budget range; whether the project is fully or under-funded; whether projects have actually begun; target populations: gender, age group, occupation and/or ethnicity; descriptions or keywords that more fully describe projects; how a project fulfils the goals in the NSP. Some specific reports are: a full report on an individual project; all organizations implementing HIV projects; HIV projects by location; projects undertaken by an executing/implementing organization; funds and technical support committed by resource provider, executing/implementing; organization or project; responses in relation to a particular target group; responses in relation to a particular keyword or type of activity; responses in relation to a particular sub-national level; responses in relation to the strategic objectives of an NSP; activities by age group or gender in relation to the NSP.

- iii. The research inventory database: The research inventory database enables countries to track research related to HIV/AIDS and sexually transmitted infections (STIs). This is a simple compilation of information on all HIV/AIDS-related research being undertaken at country level. This research mapping will facilitate identification and contact with key researchers' in-country to ensure that analysis undertaken in relation to information within CRIS is informed by local research findings.

1.4 Research problem and objective

1.4.1 Statement of the problem

The review of the M&E software landscape conducted in section 1.2 and 1.3 brings out a number of issues that help to shape the direction of this research and a hierarchy of these issues is presented below:

- i. The M&E software developed for use by specific organisations (e.g. PEPFAR, PEMS, LOGICS) exhibits a lot of inflexibility to evolving requirements. The review, for example, reveals that each of the three products was modified at least once in order to accommodate emerging requirements. The practice of re-programming the software with every change in user requirement escalates cost of maintaining software and is a constant reason for software abandonment. Yet, changes

to existing programme or project is something that even the theory of social programming acknowledges. According to Cook and Shadish (1986), projects are frequently added, modified or removed from existing programmes. Likewise, elements are also frequently added, modified or removed from existing projects. As such, a programme or project, and the related documents guiding its evaluation such as M&E framework/plan are therefore also in a constant flux. And yet many of the custom-developed M&E software reviewed above were developed on the basis of information contained in the M&E plan and framework. It is apparent therefore that a need for an alternative source of documentation for development of M&E software is necessary.

- ii. Two examples of software which are considered adaptable, CRIS and eM&E™, fall short of being truly adaptable software. CRIS was found to be adaptable (to some extent though) only within the HIV/AIDS sub-sector, and specifically in relation to the UNAIDS operations. For eM&E™, the claim that it was configurable to different M&E needs could not be substantiated due to insufficient documentation. Herein lies another problem: little effort has been made to document the requirements for adaptable MEIS. Industries such as project management, manufacturing, banking and insurance boast of several off-shelf applications just because prior efforts were directed towards developing conceptual or information models for those disciplines (Ahlemann, 2009; Fettke & Loos, 2003). Unfortunately, there is no known information model for the M&E domain. A sequential search for the keywords “reference model*”, “conceptual reference model*” and “conceptual model*” in Science Direct, Emerald and Google Scholar returned no positive results for the M&E domain.

The issues identified above portray a need for research in conceptual modelling for M&E information systems.

1.4.2 Research objective

This study is descriptive in nature and aims to develop a conceptual model for programme monitoring and evaluation information system. The specific objectives are to:

- i. Undertake a detailed review and analysis of scholarship with a view to identify the evaluation models that are available and profile their key concepts and characteristics;
- ii. Carry out critical analysis of selected M&E reports in order to identify and describe the types of data that a M&E study utilizing the key concepts identified in objective (i) typically collects and analyzes; and
- iii. Develop a conceptual model basing on the information gathered from objectives (i) and (ii).

1.5 Design and Methodology

This study is designed as a non-empirical study, and more specifically, it corresponds to a class of design type which Mouton calls “theory-building or model-building studies” (2008).

Inquiry into the research question posed in this thesis is pursued through the qualitative research paradigm. More specifically, the study adopts a descriptive approach in identifying, documenting and analysing patterns and relationships within and between concepts that were extracted from various monitoring and evaluation reports. A total of fourteen reports selected from three HIV/AIDS-related programmes implemented in Uganda are used and selection of the reports is purposeful in order to achieve representation in terms of addressing both programme monitoring and evaluation issues.

The study is guided by an analytical framework that emerges from the review of literature on evaluation models, systems thinking and conceptual modelling. The Entity, Functional Schema and Relation are the major dimensions of the analytical framework that is used to guide the data collection, analysis and conclusion.

1.6 Layout and structure of the thesis

The structure of the remaining chapters is as follows:

The next chapter reviews and discusses existing classifications of evaluation models. It assesses the extent to which the models “fit” the criteria adopted in the classification scheme and concludes with a list of key M&E concepts derived from the models discussed in the chapter.

Chapter three presents an analytical framework that has been used to guide the data collection, analysis and conclusion phase of this research. The chapter highlights how information from the previous chapters and the review of scholarships on systems thinking and conceptual modelling helps in development of the analytical framework.

Chapter four provides a design map, which describes the main object of the research, the unit of analysis, specific measurements and observations to be made and the accompanying methodology and how it is employed in the research.

Chapter five presents the data collected during the research, along with a discussion of patterns and relationships observed in the data.

Chapter six discusses and synthesizes the outputs of the previous chapters, particularly chapters 2, 3 and 5. The chapter presents the final output of this study: a conceptual model for a programme monitoring and evaluation information system.

A reflection on the research journey and the various outputs generated during the research process is done in chapter seven with a conclusion that synthesizes the outputs into statements that provide recommendations for practice and further research.

Chapter 2: LITERATURE

2.1 Introduction

Evaluation is an enterprise in which measurement is central. However, the precise nature of what is measured varies from one evaluation approach to another. This thesis posits that it is possible to get a general idea about the kinds of data used in each measurement by looking at the key concepts that each evaluation approach supports. In practical terms, one would have to identify all the evaluation approaches that are available “outside there” in order to profile their key concepts. In this chapter, existing classifications of evaluation approaches are reviewed and key concepts that are characteristic of each approach identified. The review sets the ground work for the analytical framework developed in Chapter 3.

2.2 Classification of evaluation models

The practice of evaluation, while is grounded in traditional social science approaches, is quite complex (Weiss, 2005). Evaluation studies require empirical evidence to be integrated with decisions on standards and values to reach robust evaluative conclusions. In addition, every evaluation situation is distinct, and needs tailoring to suit the purpose; the evaluator’s preference of approaches and the nature of the evaluator-stakeholder relationship (Rossi, et al, 2004; Weiss 2005). This position is exemplified succinctly by Weiss:

if our priority is making sure our audiences use our work, we might choose a utilization focused approach. If our priority is answering as unequivocally as possible “what works,” we may choose a randomized trial. If our priority is engaging stakeholders and building evaluation capacity, we may choose an empowerment or participatory approach (2005:1).

Over the years, theorists have developed a wide range of models of evaluation practice based on diverse beliefs about how evaluation ought to be organized and conducted. The result is a proliferation of models, which present practitioners with a selection dilemma (Hansen, 2005). There have been attempts to collapse the various approaches into a few basic “schools” or “traditions”, although there has not been a generally accepted criterion for such a classification (Vedung, 1997). In this chapter, the classifications of evaluation models by Vedung (1997) and Stufflebeam (2001) is discussed.

This discussion is structured around a framework comprising of six descriptors, four¹¹ of which were adopted from Stufflebeam (2001) and two¹² being the author’s own initiative. These descriptors have

¹¹ Organizer; purpose; question; method.

been selected because they discuss the fundamental or structure of an evaluation approach - the tangible or intangible attributes of the approach. An overview of the six descriptors is provided below:

Table 1: Evaluation descriptors

DESCRIPTION	:	Provides a brief summary of the evaluation approach being discussed;
METHOD	:	Provides information on tools, techniques and procedures that are employed in conducting the evaluation;
ORGANIZER	:	Provides information on the main cue that is used in setting up the evaluation;
STAKEHOLDER	:	Provides information on the stakeholders who are involved in the evaluation process, and the nature of their involvement;
PURPOSE	:	Provides information on why the evaluation is conducted;
QUESTION	:	Provides information on the kinds of questions that are addressed in the evaluation study;

Source: Adapted from Stufflebeam, 2001

2.3 Evert Vedung

The classification by Vedung (1997) is influenced by his view and beliefs about evaluation. He defines evaluation as *careful retrospective assessment of the merit, worth, and value of administration, output, and outcome of government interventions, which is intended to play a role in future practical situations*. The definition is aligned with his desire to focus evaluation on satisfying the demands of public service and governmental affairs; although he acknowledges that the target of evaluation is much wider than just public policies and programmes. He disregards ex ante studies such as needs assessment, forethought evaluation/analysis - arguing that they are not proper evaluation. In other words, his definition excludes all studies that scrutinize courses of action that are only considered on paper but not yet adopted even as prototypes.

In keeping with his definition and beliefs about evaluation, he developed a taxonomy of eleven (11) evaluation models. The taxonomy is based on the evaluation organizer and groups the eleven models according to:

¹² Description; stakeholder.

- Effectiveness models: models that focus on the results of a given performance, programme or organization. There are seven (7) evaluation models classified under this category;
- Economic models: models that aim to relate assessment of results to the input used. There are three (3) evaluation models classified under this category;
- Professional model: a model that focuses on the subject matter only indirectly, with the immediate focus put on the question of who should perform the evaluation. The Peer-review process is provided as an example of the professional model and is the only one that Vedung identifies as relevant for evaluating public policy and programme. But the Peer-review model does not have its origins in Programme Evaluation studies, but in Research evaluation studies. And who would constitute the “peer” in programme evaluation anyway? For this reason, the professional model is considered to be unrelated to the thesis and is therefore omitted from further discussion.

2.3.1 Effectiveness model

Vedung (1997) describes evaluation approaches belonging to the effectiveness category as studies that are founded on a desire to assess the results of a particular policy or programme. He classifies the following seven evaluation approaches under the effectiveness category: goal-attainment model; side-effect model; goal-free evaluation model; comprehensive evaluation model; client-oriented model; stakeholder model - North America and policy commissions (Sweden). This section gives a brief discussion about these seven evaluation approaches along with the six descriptors listed in the previous section (description, method, evaluation organizer, stakeholder, purpose and question).

Description

Vedung (1997) describes the goal-attainment model as an approach to evaluation in which the assessment of programme effectiveness is based entirely on the goals that the programme sets itself to achieve. In this approach, the evaluator directs his efforts on assessing only results that have a linkage to the stated programme goals. The side-effect approach is much similar to the goal-attainment approach except that the evaluator must also look for programme side-effects. Therefore, the side-effect model is an approach to evaluation where the assessment of programme effectiveness is based on both the goals that the programme sets itself to achieve and all side-effects that may also result from pursuit of the stated goal. In side-effect approach, the evaluator still gears his efforts on assessing only results that have a linkage to the stated programme goals, but in addition also assesses whether side-effects that are known to associate with the stated programme goals were also produced. The goal-free approach is the opposite of the goal-attainment and the side-effect approaches. According to Vedung, the goal-free evaluation approach completely disregards programme goals and instead bases the assessment of programme effectiveness on the actual

results that the programme has produced. In the goal-free approach, the evaluator just goes on looking for any result that the programme has produced and contrasting the results with the needs of programme beneficiaries in order to make evaluative judgement on programme effectiveness. The comprehensive approach is broader in scope, but is also contingent on programme goal in its assessment of programme effectiveness. According to Vedung, the approach bases the assessment of programme effectiveness on the programme's constituent parts such as planning, implementation and results. Each programme component is assessed with a view to determine the extent to which it fosters attainment of stated programme goal(s). The client-oriented approach bases the assessment of programme effectiveness on the goals, expectations, concerns, desires, values, assumptions or needs of one category of programme stakeholders - the programme client or target beneficiary. The evaluation only progresses on the basis of the information that the programme clients want the evaluation to answer or seek. The stakeholder-model (North America) is similar to the client-oriented approach except that involves a broad range of stakeholders in the evaluation process. As described by Vedung, the stakeholder (North America) approach bases the assessment of programme effectiveness on the concerns and issues of all the people who have an interest in or are affected by the programme - in other words all the programme stakeholders. The evaluation only progresses on the basis of what each stakeholder group wants to know about the programme however diverse the issues might be. The Swedish version of the stakeholder-model, what Vedung also called ad hoc policy commission bases the assessment of programme effectiveness on the concerns and issues of all the people who have an interest in or are affected by the programme - in other words all the programme stakeholders. However, the evaluation is conducted by an ad hoc policy commission that has representation from the various stakeholder groups. In this approach, evaluation of the programme or policy is just one of several inputs into the process of formulating a new policy option.

Method

Effectiveness evaluation approaches all aim to measure effectiveness of a programme or policy. However, they do so using a variety of approaches, techniques and methods. This subsection describes the methods used by each of the seven approaches, beginning with the goal-attainment approach.

Conducting a goal-attainment evaluation appears to be contingent on a three-stage process. In the initial stage, the evaluator attempts to understand and make sense of the programme goals. He clarifies the programme goal and builds a common understanding of it among those with interest in the evaluation. The second stage of the process involves the collection of facts that may show the extent to which the programme goals have been realized. The result of this second phase is a comparison between the levels of achievement planned for each goal and the actual results. The final stage of the process is concerned primarily with assessing whether the observed results was actually

produced or caused by the programme or policy. Vedung goes ahead to suggest that evaluation studies that proceed only up to the second stage should be called “goal-attainment measurement” while those that advance up to the third stage should be called “impact assessment” studies.

The side-effect evaluation approach follows similar procedures to those described under the goal-attainment approach except that in the first stage, there is an additional activity of determining what the criteria of side-effects are. The evaluation then proceeds to the second and / or third stage in similar manner to what has already been described under the goal-attainment approach.

The goal-free approach starts its methodology with the evaluator uncovering all the effects that a programme has produced without any consideration to what the pre-mediated goals of the programme are. At the initial stage, the evaluator concentrates on understanding what the programme is doing and what its effects have been as opposed to what the programme should be doing. In the subsequent steps, the needs of the impacted population are contrasted with the effects uncovered in the first stage in order to determine the merit and worth of the programme. This last stage of the process is where Vedung proposes a variation to the main stream. In his view, evaluators should concentrate on presenting the bare facts and leave programme implementers and power wielders the task of assigning merit and worth.

The comprehensive evaluation approach is conducted on the basis of pre-defined criterion of merit and worth that is tied to each component of the programme. The approach recognizes the following three components of a programme: (1) the planning phase; (2) implementation phase; and (3) outcome phase. For each component, the evaluator provides a description of what the programme intended to achieve against what was actually achieved. This description is guided to a great extent by the pre-defined criterion for the component being evaluated. The analysis of intent versus actual provides the evaluator with a rich information base for passing judgement on the performance of the programme.

The client-oriented approach starts its methodology with the identification and selection of a sample of the programme beneficiary or target group. Once the programme beneficiary has been identified, a next step of the process seeks to determine what the programme clients would like to know about the programme. The needs and desires of the programme clients become a basis by which the evaluation is organised and conducted. Programme clients may also be asked to pass judgments on some aspects of the service, including asking clients to estimate programme impact. Such strategy may ask the clients to compare what would have happened had there been no programme to what actually happened with the programme in place. It is possible for clients to have diverse views of the programme effect or value of the programme. This approach accepts divergent perspectives and acknowledges that consumers can disagree on their assessment of the programme.

The stakeholder (North America) approach begins with a mapping of who the major programme stakeholders¹³ are and what their issues¹⁴ and concerns¹⁵ are. The issues and concerns of the programme stakeholders become the basis by which the evaluation is organised and conducted. At reporting; the evaluator prepares separate reports for each stakeholder group. The report discusses only the peculiar issues and concerns raised by the individual stakeholder groups. The responsibility of passing judgement on programme effectiveness is vested with each stakeholder group although the evaluator collates the individual assessments into a final report.

The stakeholder (Sweden) evaluation approach starts with the constitution of an ad hoc policy commission that has representation from the major stakeholder groups. The commission's initial activity is commissioning of expert-studies such as evaluation of past or current policies and programmes. The results of the expert-studies are collated into a proposal, which is used as a basis for soliciting inputs/comments from various stakeholders within a fixed timeline. These responses, together with the ad hoc commission proposal are debated in the public sphere resulting into the preparation of a draft bill that is tabled to parliament for approval.

Evaluation organizer

An evaluation organizer describes the main cue that is used in setting up evaluation of a particular type. Simply put, all the organizers required for conducting a particular type of evaluation study must first be in place before the evaluation can proceed. For the goal-attainment evaluation, the pre-mediated programme goals are used as the main cue for planning and conducting the evaluation. As a result, less time is spent locating the evaluation organizer since programme goals are usually contained within the programme design documents and programme implementers have a good knowledge of them. The side-effect evaluation approach uses both the programme goal and side effects as organizers of the evaluation. Therefore, some preliminary studies must be carried out to locate side-effects that are relevant to the particular programme. This may include preliminary studies around applicable social science theories, position/apprehension held by those who opposed the programme and public controversies linked to enactment of the programme. For the goal-free evaluation approach, results of the programme are used as evaluation organizer. Just like with the side-effect approach, preliminary studies are required to profile all the results that a programme has produced before the evaluation is conducted. The comprehensive evaluation approach uses both

¹³ A listing of stakeholders applicable to public programmes adopted from Vedung(1997): Citizenry; Decision-makers; Political Oppositions; National Agency Managers; Program Directors; Regional Agency Managers; Private Intermediaries; Local Agency; Street-level Bureaucrats; Clients; Neighboring Agencies; Program Competitors; Contextual Stakeholders and Research Community.

¹⁴ Issue is any statement, proposition, or focus that allows for the presentation of different points of views; any proposition about which reasonable persons may disagree; or any point of contention (Guba and Lincoln, 1981:33ff. cited in Vedung (1997).

¹⁵ Concern is any matter of interest or importance to one or more parties. It may be something that threatens them, something they fear might lead to undesirable consequences for them or something they are anxious to substantiate.

system components and programme goals as organizers of the evaluation. Both organizers; the programme goals and the systems components are easily identifiable and as such no preliminary study is necessary. The client-oriented approach uses a multiplicity of organizers such as the expressed desires, expectations, values, assumptions, and objectives of the clients or programme addressees. This approach requires preliminary studies to elicit the desires, expectations, values, assumptions, and objectives of the clients or programme addressees. Once information on the organizers has been finalized, the evaluation proper begins. Both the North America and the Sweden stakeholder-models use the expressed desires, expectations, values, assumptions, and objectives of the various programme stakeholders as organizer of the evaluation. These organizers are similar to those specified under the client-oriented approach. The only variation is the fact that in the stakeholder models, a broad range of stakeholders are consulted unlike in the client-oriented model where only one category of stakeholder is consulted.

Stakeholder

The stakeholder descriptor provides information on the category of people who are involved in the evaluation process, and the nature of their involvement. The description that Vedung provides on the goal-attainment, the side-effect and the comprehensive evaluation approaches does not indicate that any stakeholders are involved in the evaluation exercise. It is likely that the evaluator performs the entire evaluation without any stakeholder involvement. The goal-free, client-oriented, stakeholder (North America) and stakeholder (Sweden) approaches all grant stakeholders some roles in the evaluation process. The goal-free and the client-oriented evaluation approaches grant the impacted population the roles of determining the issues that the evaluation should focus on and in some instances to attach values/worth to the evaluation findings. To this end, the evaluator consults the impacted population about their needs and desires, which constitute the criteria for merit and worth. The point of departure in the two approaches relates to the way the needs of the impacted population are put to use. In the goal-free approach, the need of the impacted population is used to attach value and worth to the results and the effects that the programme produced. On the contrary, the client-oriented approach uses the need of the impacted population to plan the evaluation. The stakeholder (North America) approach involves all the groups that have interest in or who are affected by the programme in the evaluation exercise. Involvement includes having them provide the problem to be investigated and the criteria and standards to be used in the evaluation. The stakeholder (Sweden) approach involves all groups that have interest in or who are affected by the programme but the involvement is through having representation in the ad hoc policy commission. Each stakeholder group is represented in an ad hoc policy commission, which is responsible for conducting the evaluation.

Purpose

The purpose descriptor provides a rationale for why the evaluation is being conducted. The description provided of the goal-attainment approach indicates that the evaluation is conducted for the purpose of determining whether the goals that the programme sets itself to achieve were actually achieved. This is also the same purpose for the side-effect evaluation except that the side-effect approach additionally aims to determine the positive/negative; anticipated/unanticipated effects that were produced both outside and inside of the programme's target area. The description of the goal-free approach suggests that the approach aims to determine all the impacts that a programme or policy has had on its targeted population without any consideration to what the set goals of the programme were. The client-oriented and the stakeholder (North America) approaches are based on the need to determine whether the programme satisfies the concerns, desires, and expectations of its stakeholders. The point of departure between the two approaches is in the category of stakeholders. The client-oriented approach considers only one category of stakeholder – the impacted population while the stakeholder (North America) approach considers all groups that have interest in, or are affected by the programme. The stakeholder (Sweden) approach aims to use programme evaluation to enable the formulation of policy options (or the most promising option) for future policy actions, and sometimes clear recommendation for policies.

Question

The question descriptor provides information on the kinds of issues that are addressed in the evaluation study. In the goal-attainment approach, the two main issues addressed are whether the results that the programme produces are in accordance with programme goals and whether the programme produces results. In the side-effect approach, the key questions addressed include what the anticipated and unanticipated effects of the programmes are; determining whether or not the programme achieves its set objectives and determining what the linkage between the programme and the observed effects is. For the goal-free approach, the description provided is silent on the nature of questions that the approach addresses; for the comprehensive approach, questions/issues addressed include determining whether conditions specified at the planning phase are fulfilled as specified; whether or not the programme is implemented as planned and whether the actual impacts/outcomes of the programme conform to those specified at planning. The client-oriented and the stakeholder (North America) approaches do not determine their questions before hand. Instead, they tend to address issues that point in various directions, depending on the needs and desires of the programme client or stakeholders. The stakeholder-model (Sweden) focuses on addressing the issues of developing policy alternatives and its implementation.

This concluding section brings to an end the discussion of the seven evaluation approaches characterised under the effectiveness category. The discussion was guided by the six descriptors that are developed in section 2.2. On the basis of the description, the following inferences are made:

1. The comprehensive, client-oriented and stakeholder approaches may sometimes fail to pursue results that can demonstrate programme effectiveness. The comprehensive approach may fail to measure programme effectiveness if the evaluation focuses on the programme components that do not lead to a measure of effectiveness. As Vedung noted: *“the comprehensive evaluation model can be turned into process evaluation”* (1997:65). The client-oriented and stakeholder approaches may fail to measure programme effectiveness because the organization of the evaluation is based on clients' and stakeholders' concerns, which are sometimes very diverse and un-predictable; and may not focus on programme effectiveness. For example, in the Swedish version of the stakeholder model (the Ad Hoc policy commissions); assessment of programme effects is a secondary issue. The primary focus is to assess the alternatives for future policy actions. To quote Vedung *“investigations have traditionally focused on alternatives for future action rather than impacts of past policies”* (1999:77).
2. Evaluation organizers used under the effectiveness category vary along two dimensions. On the first dimension, single or multiple organizers are used in certain types of evaluation studies. For example evaluation approaches such as the goal-attainment and the goal-free employ a single type of evaluation organizer (programme goal and programme results respectively). While evaluation approaches such as side-effect, client-oriented, comprehensive and stakeholder employ multiple types of evaluation organizer. For example, the comprehensive evaluation approach uses programme goals and systems components as evaluation organizer. On the second dimension, evaluation organizers vary along the line of whether it is intrinsic to the programme or not. Evaluation organizers used by the approaches such as goal-attainment, side-effect and the comprehensive approaches are intrinsic and are derived from within the programme. Example of such organizers includes programme objectives; goal and systems components. For this type of organizers, no special study is necessary in order to identify them. Conversely, evaluation organizers for approaches such as goal-free, client-oriented and stakeholder use organizers that are external to the programme. Such organizers include beneficiary needs, desires and assumptions. In order to identify them, the evaluator needs lengthy encounters with programme beneficiaries.
3. Involvement of stakeholders in evaluation was observed to vary along two dimensions. On the first dimension, no stakeholder, a single stakeholder or multiple stakeholders are involved in the evaluation exercise. On the second dimension, the nature of stakeholder involvement in the evaluation exercise varies to include involvement in planning (questions/issues setting; choice of methods); execution of the evaluation plan and interpretation of results.

To facilitate further inference, the characteristics of the seven effectiveness evaluation approaches along the dimensions of whether it uses single or multiple types of organizer are mapped; whether the organizers used are intrinsic or external to the programme; the various categories of stakeholders involved in the evaluation; and the role(s) stakeholder(s) play in the evaluation, if any. The outcome of the exercise is summarized in the table below.

Table 2: Characteristic of effectiveness approaches on organizer & stakeholder descriptors

Approach	Organizer (single / multiple)	Source (Intrinsic / External)	Stakeholder (None / Single / Multiple)	Role (none/ execute / interpret) plan/ interpret)
Client-oriented	Multiple	External	Single	Plan
Comprehensive	Multiple	Intrinsic	None	None
Goal-attainment	Single	Intrinsic	None	None
Goal-free	Single	External	Single	Interpret
Side-effect	Multiple	Both	None	None
Stakeholder (North America)	Multiple	External	Multiple	Plan
Stakeholder (Sweden)	Multiple	External	Multiple	Plan, Execute, Interpret

There are at least two interesting patterns that are observed from the information summarized in the above table. The first pattern shows that stakeholder participation in a given evaluation varies depending on whether the organizer used in the evaluation is intrinsic or external to the programme. Where the organizer used is external there is stakeholder involvement in the evaluation. And where the organizer used was intrinsic to the programme, there is no stakeholder involvement in the evaluation exercise. Plausible explanation for this variation can be inferred from the process that goes into the development of a programme or policy. This process is usually broad-based and consultative. A wide-range of stakeholders is involved and consensus reached on the parameters of the programme or policy such as programme goals, objectives, strategies, expected outcomes. Therefore, if the evaluation is based on these parameters for which consensus has already been reached, there might be no need for stakeholder consultation. Any engagement with them would be purely for the sake of clarifying and gaining better understanding of the parameter. On the other hand, where the organizer is external to the programme, the evaluator needs to involve stakeholders for possibly several reasons including identifying the organizer. The second pattern shows that whenever the evaluation is geared at serving the interest of multiple groups of stakeholders, there is always a

role that stakeholders play in the process. Both the North America and the Swedish versions of the stakeholder model target multiple groups of stakeholders and also involve them in planning, executing and interpreting the evaluation finding.

2.3.2 Economic model

Economic model describes evaluation approaches that assess the results of a public policy or programme in relations to the cost incurred. Models belonging to this category must emphasize both the aspects of programme results and costs. Models classified under the economic category include *productivity* and *efficiency* models.

Description

According to Vedung, productivity is an evaluation approach that attempts to measure the performance of a public intervention or programme through the assessment of the relationship between output of products and/or services and input of resources. In simple terms, the approach tries to solve the mathematical formula: output divided by input (output/input). Here, input can be operationalized either as a monetary entity or as a physical/non-monetary entity. Vedung also describes efficiency model as an evaluation approach that attempts to relate costs to programme impacts. An efficiency assessment is of two types: cost-benefit and cost-effectiveness analysis. In cost-benefit analysis, both programme inputs and outcomes are measured in monetary terms, whilst in cost-effectiveness analysis, inputs are estimated in monetary terms but outcomes in terms of actual impacts.

Method

The evaluation approaches discussed in this section all aim to relate programme effectiveness with cost incurred in the production of programme results. However, they do so using a variety of approaches, techniques and methods. For the productivity approach, the method involves determining the criterion of productivity. Methods for the productivity model start with operationalizing the formula: output/input. The next step involves setting or determining standards for knowing good and bad performance (e.g. comparison with past performance or with similar institutions in the same country or with similar institutions in other countries or with goals of the political bodies or with client/stakeholder goals). In the last step, performance data is collected and analysed to show performance. In the productivity model, the first phase of the method involves the evaluator discerning programme effects. These are effects caused by the policy or programme, and nothing else. Next, depending on whether it is a cost-benefit analysis, programme effects are converted into monetary values and compared with programme cost. Otherwise, the hard, physical impacts are compared with programme costs.

Evaluation organizer

An evaluation organizer describes the main cue that is used in setting up evaluation of a particular type. Simply put, all the organizers required for conducting a particular type of evaluation study must first be in place before the evaluation can proceed. For the productivity approach, productivity is used as the evaluation organizer. For the efficiency model, Vedung only indicates that the evaluation organizer varies depending on the sub-type selected. The sub-types include cost-benefit and cost-effectiveness evaluation.

Stakeholder

The stakeholder descriptor provides information on the category of people who are involved in the evaluation process, and the nature of their involvement. Vedung's description of the productivity and efficiency approaches is silent on whether any stakeholders are involved. In these approaches, the evaluator performs the entire evaluation without any stakeholder involvement.

Purpose

The purpose descriptor provides a rationale for why the evaluation was conducted. The productivity evaluation approach is conducted with the purpose of determining the level of performance of a public intervention/programme. The efficiency evaluation approach is conducted with the purpose of determining the effectiveness of a programme and the cost incurred.

Question

The question descriptor provides information on the kinds of issues that are addressed in the evaluation study. In both the productivity and efficiency approaches, the manuscript is silent on the typical questions that the evaluation approaches address.

This concluding section brings to an end the discussion of the two evaluation approaches characterised under the economic model category. The discussion is guided by the six descriptors that were developed in section 2.2. On the basis of the description, the following inferences are made:

1. The classification of evaluation approaches as being of type "economic" is driven by the criteria that the model assesses programme outcomes in relation to programme costs (inputs). Assessment establishes that all the models discussed under this category do actually support this criterion. However, the distinction between the productivity and the cost-effectiveness models is thin. Both models compare the cost of generating programme

outputs to cost incurred - except that cost-effectiveness looks at “big” outputs in the form of outcomes and impacts while productivity looks at instrumental outputs: outputs that result directly from execution of activities.

2. The evaluation approaches under this category use broad parameters as evaluation organizer. The evaluator is required to conduct some preliminary studies to operationalize them. The approaches under this category also do not give stakeholders any role in conducting the evaluation.

This section appraises Vedung's classification of eleven evaluation models according to his proposed categories of effectiveness; economic and professional models. Models under each category are assessed and existing overlaps identified. The eleven models appear to fit well within the suggested categories, with negligible overlap across categories - a minor exception being the close linkage between effectiveness and economic models (particularly the efficiency models). It is observed that both effectiveness and efficiency models involve the assessment of programme effects, with the efficiency model employing the additional step of relating the effects to cost. In a way, effectiveness models are a sub-set of the economic model, particularly the efficiency model.

Assessment of the evaluation approaches against evaluation organizer and stakeholder involved reveals interesting variations. In summary, if the evaluation is based on evaluation organizers for which consensus had already been reached, there is no need for stakeholder consultation. On the other hand, where the organizer used is external to the programme, and no consensus has previously been reached, there was always stakeholder involvement. Also, whenever the evaluation is geared at serving the interest of multiple groups of stakeholders, there is always a role that stakeholders play in the process.

2.4 Daniel Stufflebeam

The classification of evaluation models by Stufflebeam (2001) is driven by a desire to assess which of the 20th century approaches are valuable for future use and which ones are best discarded. He reviews twenty-two evaluation approaches according to a framework¹⁶ comprising of ten descriptors. He defines evaluation as the *assessment of an object's merit and worth*, and categorizes the twenty-two approaches into four broad groups:

- i. Pseudo evaluation approaches – representing politically motivated types of evaluation in which evaluators are tempted to shade, selectively release or falsify findings in order to fulfil a political interest. This is done so that a positive or negative view of the programme is

¹⁶ Advance organizer; purpose; source of questions; characteristic questions; methods commonly employed; theorists; when to use; strength of the approach; weakness of the approach

portrayed, contrary to its true image. This category does not represent “true” evaluation approaches and is therefore not discussed further;

- ii. Question/method-oriented approaches - representing types of evaluation that are driven by the need to answer certain questions or a desire to apply certain methodologies. The issue of whether the questions or methodologies lead to the central issue of measuring a programme's merit and worth is secondary. There were thirteen approaches discussed under this category;
- iii. Improvement/accountability approaches - representing types of evaluation studies that are designed primarily to demonstrate accountability or to provide information that may be used to improve a current or future programme. Approaches under this categorization employ a broad spectrum of value criteria to assess all the parameters relevant to measure a programme's merit and worth. Three approaches are discussed under this category;
- iv. Social agenda/advocacy approaches - representing types of evaluation that are oriented towards employing the perspectives of stakeholders as well as experts in characterizing, investigating, and judging programs. The approaches emphasize the importance of democratically engaging stakeholders in obtaining and interpreting findings. They also stress serving the interests of underprivileged groups. There are four approaches discussed under this category.

2.4.1 Question/Method-Oriented Studies

The question/method-oriented studies correspond to evaluation activities that are geared towards answering a set of pointed questions or towards employing specific methods or methodologies. Consequently, the evaluation is less concerned about whether the questions or methods can provide sufficient evidence to measure a programme's merit and/or worth. For this reason, Stufflebeam (2001) has sometimes referred to these approaches as quasi-evaluation because they may only partially measure a programme's merit, worth or significance. Stufflebeam identifies thirteen approaches under this category: objectives-based studies; accountability (particularly payment by results); objective testing; outcome monitoring/value-added assessment; performance testing; experimental studies; management information systems; benefit-cost analysis; clarification hearing; case study evaluations; criticism and connoisseurship; programme theory-based evaluation and mixed methods studies. However, the objective testing; outcome monitoring/value-added assessment and performance testing approaches have been excluded from this summary because they are found to be specific to only educational programmes (Stufflebeam, 2001). This makes them inappropriate as sources of information for developing a conceptual model that is expected to be generic to all evaluation approaches. In the subsequent sections, the question/method-oriented approaches are discussed along the following six descriptors: description, method, evaluation organizer, stakeholder,

purpose and question. The discussion is a summary of the descriptions that Stufflebeam (2001) provides in his manuscript on these evaluation approaches.

Description

This section provides a summary description of the ten evaluation approaches classified as questions/methods oriented. The description is a summary of the discussion Stufflebeam (2001) advances on each of the approach.

Daniel Stufflebeam describes objective-based approach to evaluation as a set of evaluation activities that aim to assess whether a programme has achieved its set objectives. He indicates that the evaluation is conducted by internal evaluators and is performed after the programme has terminated or ended. He describes the accountability (particularly payment by results) evaluation as an approach that focuses on ascertaining whether a programme produced the desired outcomes. Unlike the objective-based approach, the accountability study is conducted by an external evaluator and is guided by a set of pass/fail standards. Depending on what the established pass mark is, payments are made for good results, while sanctions are given for unacceptable performance. Assessment is always an ongoing and continuous exercise and not merely a once-off activity. He describes the experimental evaluation as an approach to evaluation that uses the experimental research technique such as the experimental and control groups to determine un-equivocally whether observed programme outcomes were caused or a result of the programme. The description provided of the approach suggests that assessment is performed on an ongoing programme, although its variances such as quasi-experimental studies could be applied to completed programmes. The Management Information Systems approach to evaluation focuses on the identification, collection, storage, analysis and dissemination of information that programme managers could use to plan, direct, control, and report on matters within their spheres of responsibility. The discussion of the approach suggests that evaluation is carried out on an ongoing basis and on a living programme or project, and is performed by internal staff/evaluator. The cost-benefit analysis approach describes evaluation activities that are conducted with the aim of determining the cost incurred in achieving programme objectives. Costs associated with programme inputs, outputs and outcomes are collected, analyzed and used to pass judgement on the merit, worth or significance of the programme. The Clarification hearing approach denotes evaluation activities that aim to measure programme effectiveness, but with the evaluation conducted along the judicial system. In the approach, evaluators become 'lawyers' and using agreed-upon criteria of programme effectiveness, argue in defence or against programme achievements. Passing of final judgement on whether the programme was effectiveness or not was performed by an evaluator playing the 'judge' role. The case study evaluation approach denotes evaluation activities that are conducted with the aim of providing an authoritative, in-depth and well-documented description of a programme. The approach gives no primacy to the issue of programme effectiveness

and instead prefers to focus on providing in-depth description, analysis, and synthesis of a particular programme. The Criticism and Connoisseurship approach describes evaluation activities conducted to describe, critically appraise, and illuminate the merits of a programme. However, the approach dictates on who should conduct the evaluation. This is because the approach assumes that certain experts in a given substantive area are the ones capable of in-depth analysis and evaluation that could not be done in any other ways. The programme theory-based approach denotes evaluation activities that are conducted with the aim of understanding how the theories and mechanisms undergirding a particular programme fostered production of programme outcomes. The approach employs the theory or mechanism by which programme activities are understood to produce outcomes for designing the evaluation study and for interpreting its findings. The mixed-method approach to evaluation denotes activities that must employ a mix of both quantitative and qualitative methods to conduct a given evaluation study. The approach is preoccupied with using multiple methods and therefore accords little priority to the measurement of merit and/or worth. Both summative and formative assessments are supported, either to provide direction for improving programmes or to assess their effectiveness after they have had time to produce results.

Method

The question/method-oriented approaches denote evaluation activities that are geared towards answering a set of pointed questions or towards employing specific methods or methodologies. In the objective-based approach, evaluation appears to follow a two-step process. In the initial stage the evaluator attempts to understand and make sense of the programme objectives by clarifying and describing them in clear and measurable statements. The second stage of the process involves the collection and analysis of facts to determine how well each programme objective has been realized. Stufflebeam observes that the objective-based approach also draws a lot of inspiration from management-oriented approaches such as management by objective and a wide-range of performance oriented assessments. Stufflebeam describes the methodology for the accountability approach as revolving around two strands: (1) procedures for setting pass/fail standards upon which a good performance could be differentiated from a bad performance; and (2) procedures for collecting and analysing information that could distinguish a good or bad performance. He lists management-oriented approaches such as performance contracting and management by objectives as methods that could be used. Although not explicit there is an inclination to think that the management-oriented approaches are used in setting of pass/fail standards. Procedures for collecting and analysing information include programme input, institutional report cards/profiles, process and output databases, programme planning and budgeting systems. Collection of information is performed on a continuous basis using variety of methods such as self-studies, peer reviews; mandated testing programs. Methods for the experimental approach are based on: (1) identification and random assignment of programme beneficiaries to experimental and control groups; and (2) contrasting

outcomes of the experimental group with those of the control group after the experimental group receives a particular intervention and the control group receives no special treatment or some different treatment. Statistical methods such as study of outliers and cross-break tables are then used to draw evaluative conclusions. The management information system evaluation supports its method around two strands: in the first strand, the evaluator employs a combination of techniques such as systems analysis; programme evaluation and review techniques; critical path method; management by objectives; and programme planning and budgeting system to identify the information that programme management personnel and their superiors need in order to plan, direct, control, and report on matters within their spheres of responsibility. The second strand involves continuous collection of information identified in the first strand. According to Stufflebeam, the techniques for the management information system approach include computer-based information systems; periodic staff progress reports; regular budgetary reporting and input process and output database. Methods for the benefit-cost evaluation approaches are inclined to revolve around the analysis of costs. Procedurally, the approach breaks down cost analysis into three levels: (1) cost analysis of programme inputs, (2) cost-effectiveness analysis, which focuses at computing the costs associated with achieving each objective or outcome; and (3) benefit-cost analysis, which focuses at computing the costs associated with achieving each objective or outcome and the monetary value of the outcomes. The clarification hearing approach uses methods similar to the judiciary system and its trial procedures and methods. The major tenets of the method include trial proceedings, expert critics, hearings, forums and testimony. In the initial stage agreement is sought on the trial procedures as well as the criteria to use for measuring programme effectiveness. In the next phase, evaluators acting as “lawyers” employ the agreed-upon criteria of programme effectiveness to argue in defence or against programme achievements. Basing on the evidence provided by both sides a judge passes final judgement on whether or not the programme was a success. The case study approach is open to the use of both qualitative and quantitative approaches. The approach only cares for whether the selected methods allows for the discovery, categorization and in-depth / well-documented description of the programme. Examples of methods for the case study evaluation include analysis of archives; collection of artefacts, such as work samples; content analysis of programme documents; both independent and participant observations; interviews; logical analysis of operations; focus groups; tests; questionnaires; rating scales; hearings; forums; and maintenance of a programme database. The criticism and connoisseurship approach employs a more tacit perspective on how evaluation ought to be conducted. The evaluator, being the expert employs perceptual sensitivities, past experiences, refined insights to conduct the evaluation. The evaluator’s judgments are conveyed in vivid terms to help the audience appreciate and understand the programme’s entire outlook. The programme theory evaluation approach bases its method on a model of the programme’s logic. The way in which the programme is perceived to operate and to produce required outcomes is described in the form of a detailed flowchart. The detailed flow-chart is then used to guide the selection of

questions, indicators and assumed linkages between and among programme elements that should be used in evaluating the programme. Like the case study approach, the mixed method evaluation approach is open to the use of both qualitative and quantitative approaches. No single method is prescribed and all that matters is selecting methods that can generate dependable feedback on a wide range of questions; depth of understanding of particular programme; a holistic perspective; and enhancement of the validity, reliability, and usefulness of the full set of findings. Examples of methods for “mixing” include surveys using representative samples, cohort and cross-sectional samples, norm-referenced tests, rating scales, quasi experiments, significance tests for main effects, ethnography, document analysis, narrative analysis, purposive samples, single cases, participant observers, independent observers, key informants, advisory committees, structured and unstructured interviews, focus groups, case studies, study of outliers, diaries, logic models, grounded theory development, flow charts, decision trees, matrices, and performance assessments.

Evaluation organizer

The evaluation approaches discussed under the question/method oriented category employ various types of organizers for setting up the evaluation study. The objective-based evaluation approach employs the programme objectives as organizer for the evaluation study. In effect, there is less time spent locating the evaluation organizer since programme objectives are usually contained within the programme design documents and programme implementers have a good knowledge of them. Both the accountability evaluation and the experimental evaluation approaches employ a multiplicity of evaluation organizers. The evaluation organizers for the accountability approach include the persons and groups responsible for producing results, the service providers’ work responsibilities, the expected outcomes, pass/fail cut scores, and defined consequences of passing or failing. For the experimental approach, evaluation organizers include problem statements, competing treatments, hypotheses, investigatory questions, and randomized treatment and comparison groups. For both approaches, the initial phase of the evaluation involves uncovering the nitty-gritty of each of the specified evaluation organizer. The Management Information System approach employs programme objectives, specified programme activities, projected programme milestones or events and programme budget as evaluation organizers. The initial phase of the evaluation exercise involves uncovering the nitty-gritty of each of the specified evaluation organizers. The benefit-cost analysis approach also employs a multiplicity of evaluation organizers under each of the three categories. Under the cost-analysis category, the following organizers are specified: cost breakdowns for both programme inputs and outputs by line items. Under cost-effectiveness, the costs associated with programme objective are used as the evaluation organizer. For the benefit-cost analysis, the organizers that are specified include costs associated with main effects and side effects, tangible and intangible outcomes, positive and negative outcomes, and short-term and long-term outcomes. The initial phase of the evaluation is expected to focus on gathering information on each of the specified

evaluation organizers. The clarification hearing approach employs the criteria of programme effectiveness as evaluation organizer. During the initial phase of the assessment, the evaluator is expected to profile what the criteria of programme effectiveness were before the evaluation proper can begin. The case study approach employs the following evaluation organizers: definition/rationale of the programme, characterization of its geographic and organizational environment, the historical period in which it was to be examined, the programme's beneficiaries and their assessed needs, the programme's underlying logic of operation and productivity, and the employees' roles and responsibilities. During the initial phase of the assessment, the evaluator is expected to profile the nitty-gritty of the evaluation organizers. The criticism and connoisseurship utilizes the evaluator's special expertise and sensitivities as the evaluation organizer. These parameters are tacit to the expert and vary from one expert to the other. The programme theory-based evaluation approach employs a multiplicity of evaluation organizer including mechanisms by which programme activities are understood to produce or contribute to programme outcomes, along with the appropriate description of context, specification of independent and dependent variables, and portrayal of key linkages. During the initial phase of the assessment, the evaluator was expected to profile the nitty-gritty of these evaluation organizers. The mixed methods evaluation approach uses a broad-spectrum of evaluation organizers including formative and summative evaluations, qualitative and quantitative methods, and intra-case or cross-case analysis. There is a clear indication that the organizers for the mixed method approach include many of the methods in the other evaluation approaches discussed above and the evaluator is therefore required to be knowledgeable in a wide range of evaluation approaches to be able to successfully conduct a mixed method evaluation.

Stakeholder

This section describes the nature of stakeholder involvement, if any in conducting a particular type of evaluation. For the objective-based approach there is not explicit indication that any stakeholder is involved in conducting the evaluation. Stufflebeam only indicates that the audience for the objective-based evaluation studies includes programme developers, sponsors, and managers and that the questions for the evaluation can be mandated by the client, formulated by the evaluator, or specified by the service providers. Therefore, the objective-based approach is considered to be conducted with no stakeholder involvement. For the MIS approach programme, management personnel and their superiors are involved in planning the evaluation with the evaluator required to consult this group of stakeholders during the initial planning or preparatory phase of the evaluation so that their information needs are captured accurately. The case study approach involves a broad spectrum of stakeholders in the planning and conduct of the evaluation. Stakeholders that are involved include the programme oversight body, administrators, staff, financial sponsors, beneficiaries, and potential adopters of the programme. These stakeholders constantly interact with the evaluator so that the evaluation is geared towards explicating information and issues that are of great interest to them. Stakeholders

may also be called upon to interpret the evaluation findings. The Mixed method approach likewise involves a broad spectrum of stakeholders including programme administrators and staff, policy boards, financial sponsors, beneficiaries, taxpayers, and programme area experts. They are mainly involved in identifying the evaluation questions. In the Clarification Hearing approach, stakeholders are involved through the jury. The jury is composed of persons representing the various stakeholder groups. A jury hears the proceedings and issues a ruling based on the strength of the evidence presented by both sides. The accountability; benefit-cost analysis; criticism and connoisseurship; experimental; and programme theory approaches do not demonstrate stakeholder involvement in conducting the evaluation studies and prefer to rely on the expertise of the evaluator to plan, conduct, synthesize and report on the evaluation result.

Purpose

The purpose descriptor provides a rationale for conducting the evaluation. The objective-based evaluation approach is mainly concerned with issues relating to whether or not the programme achieves its objectives. The accountability approach addresses a multiplicity of issues ranging from providing relevant stakeholders with an accurate accounting of results; influencing the attainment of positive outcomes or results through the use of strategies that instil fear among programme administrators; pinpointing responsibility for good and bad outcomes; providing information bases that policy makers can use to issue policy decisions and ensure that there is standardization of outcome measures across the board. The experimental approach is mainly concerned with issues of whether or not causal relationships exist between independent and dependent variables. Like the accountability approach, the experimental approach provides information base used for issuing policy decisions in addition to comparing performance of competing programmes. The purpose for conducting management information system evaluation includes providing programme managers with continuous information for decisions relating to planning, steering and controlling; determining the extent to which programme objectives are being realized; providing information on programme shortcomings; providing direction for programme improvement; informing management decisions and actions. The benefit cost analysis approach is conducted to gain clear knowledge of the resources invested and the returns on the investment in terms of results and the monetary value of the results. The clarification hearing approach is conducted to gain better insight into the programmes' strengths and weaknesses and what results are produced by the programme. The case study evaluation approach is concerned with issues relating to provision of authoritative, in-depth and well-documented explication of the programme. The criticism and connoisseurship approach is mainly concerned with issues of appraising and illuminating the merits and shortcomings of a particular programme. The programme theory approach is concerned with providing information that can explain theoretical assumption undergirding a particular programme; whether the theories worked or did not work and why or why not. The mixed method approach is concerned with two main issues:

providing direction for improving programmes as they evolve and assessing effectiveness of a particular programme.

Question

The question descriptor provides information on the kinds of issues that are addressed in the evaluation study. In the objective-based study, questions of concerns include the extent to which stated programme objectives are achieved and whether performance meets or exceeds set standards. The accountability approach is concerned with addressing questions such as whether programme implementers are carrying out their duties and delivering results according to plans; whether current performance is better or worse than past performances and whether the programme is worthy of continuation or dissemination. For the experimental approach the questions of concern include what effect the programme has on specified outcome variables; how the programme fares when compared to other programmes addressing similar or same objectives; the effect that the programme produces and whether the programme is worthy of continuation or dissemination. The management information systems approach is concerned with questions such as whether or not activities are being implemented and results produced as per plan; whether the programme is sustainable and could be transportable to other sites; and what changes may be required in the programme design or implementation arrangements for it to produce better results. The benefit cost analysis approach is concerned with questions about return on investment; worthiness of programme for continuation and/or dissemination; and the overall cost of the programme. The clarification hearing approach addresses questions such as whether the programme is successful or not and whether it is worthy of continuation or dissemination; what changes may be required in the programme design or implementation arrangements for it to produce better results; how the programme is compared with other programmes that address similar or same objectives. The case study approach addresses questions relating to the programme in both concept and practice; how it evolves; its internal mechanism or operation; its results; its effects (negative, positive, side-effects); its short falls; its values in the eyes of its clients / beneficiaries; explanations for its failure or success; its costs; its important characteristics; what changes are required for better outcomes/results; how it compares to its competitors; what parts could be transferrable to different sites/settings; and what issues remain un-resolved. The criticism and connoisseurship approach is concerned with addressing questions relating to what the programme's core and most important characteristics are; what makes the programme different from other programmes of the same kind; what changes are required for better outcomes/results; what features can be successfully transported to different sites/settings. The programme theory approach addresses questions such as the extent to which stated programme objectives are achieved; how the programme affects and/or produces observed outcomes; what changes are required for better outcomes/results; explanation for its success or failure; and validity of hypothesized linkages and theory. The mixed method approach is concerned with addressing

questions relating to the extent to which stated programme objectives are achieved; sustainability/transportability of the programme; the effects of the programme (negative; positive; main effects; side effects); the programme's value in the eyes of its clients/beneficiaries; and what changes are required for better outcomes/results.

This concluding section brings to an end the discussion of the ten evaluation approaches characterised under the question/method category. The discussion is guided by the six descriptors that are developed in section 2.2. On the basis of the descriptions, the following inferences are made:

1. Evaluation approaches under the question/method category show conformity to the criteria with varying degrees. The objective-based and accountability approaches appear to bend more towards the "question" criteria. This implies that existence of clearly defined questions is the only thing that matters for the evaluation to proceed. The evaluator is not constrained to use any particular method or approach in answering those questions. The case study, clarification hearing, criticism and connoisseurship, mixed method, programme theory-based approaches appear to tilt more towards the "method" criteria. These approaches tend to focus attention on describing the procedures for planning, organizing and conducting particular types of evaluation studies. In essence these approaches can be used to address myriad kinds of evaluation questions. Although the criticism and connoisseurship is included as tilting towards the "method" criteria it is quite unique from the rest of the other methods. There appears to be no any substantive "method" that the approach uses besides just choosing who should perform the evaluation (the connoisseur). The chosen connoisseur then uses whatever method he sees befitting – after all, he is the undisputed expert! It is not clear whether the criticism and connoisseurship should be categorized under the question/method category as its criteria is more about who should do the evaluation than what question or method to use. The benefit-cost analysis, experimental and management information system are approaches that appear to skew towards both the "question" and the "method" criteria. A mention of the method presupposes that the evaluation will be addressing some particular kind of questions. For example, the experimental approach pre-supposes questions of cause-effect or co-relational nature. An attempt to use such methods for questions outside of their jurisdiction is likely to result into fierce criticisms.
2. Evaluation organizers under the question/method category emerge to vary along two major dimensions. On the first dimension a single or multiple types of organizer may be employed to conduct a given evaluation study. Consequently some evaluations use a single type of organizer while other evaluations use multiple organizers. On the second dimension, evaluation organizers appear to vary from being intrinsic to the evaluand to being external to the evaluand. Organizers such as programme objective, problem statements and programme mechanisms are considered intrinsic to the programme while organizers such as evaluator's

sensitivity and expertise (the organizer used in the criticism and connoisseurship approach) are considered external to the evaluand.

3. Observed variation exists in the manner in which stakeholders are involved or not involved in the evaluation. There is stakeholder involvement in some approaches while there is no stakeholder involvement in others. It is also observed that where stakeholders are involved their role varies and includes planning (determining the issues/questions that the evaluation should address), executing the evaluation and interpreting the evaluation results.

To facilitate further inference, the characteristics of the ten question/method oriented approaches are mapped along the dimensions of whether it uses single or multiple types of organizer; whether the organizers used are intrinsic or external to the programme; the various categories of stakeholders involved in the evaluation; and the role(s) stakeholder(s) play in the evaluation, if any. The outcome of the exercise is summarized in the table below.

Table 3: Characteristic of question/method approaches on organizer & stakeholder descriptors

Approach	Organizer (single / multiple)	Source (Intrinsic / External)	Stakeholder (None / Single / Multiple)	Role (none/ plan/ execute / interpret)
Accountability	Multiple	Intrinsic	None	None
Benefit-Cost analysis	Multiple	Intrinsic	None	None
Case study	Multiple	Both	Multiple	Plan/Execute/ Interpret
Clarification Hearing	Single	External	Multiple	Interpret
Criticism and Connoisseurship	N/A	External	None	None
Experimental	Multiple	Intrinsic	None	None
Management Information System	Multiple	Both	Multiple	Plan
Mixed Method	Multiple	Both	Multiple	Plan
Objective-based	Single	Intrinsic	None	None
Programme Theory-based	Multiple	Intrinsic	None	None

There are at least two interesting patterns that are observed from the information summarized in the above table. The first pattern shows that stakeholder participation in a given evaluation varies depending on whether the organizer used in the evaluation is intrinsic or external to the programme. Where the organizer used is external or both external and intrinsic there is likely to be stakeholder

involvement in the evaluation. Although from the above table the Criticism and Connoisseurship approach does not involve stakeholders even though it uses organizers that are external. It is tempting to think that this is a result of the methodology constraint placed on the approach. The approach grants maximum respect to the chosen connoisseur who is free to involve or not to involve any stakeholders in the evaluation and it is also observed that the approaches that do not involve stakeholders: accountability, benefit-cost, Clarification Hearing, Experimental, Objective-based and Programme-theory appear to be contingent on questions or issues that are less contentious among programme stakeholders either because they have been previously agreed upon or because they are obvious, fair or based on best-practice or wide-acceptance.

2.4.2 Improvement/Accountability-Oriented Studies

Improvement/Accountability approaches are designed to demonstrate accountability or to provide information that may be used to improve a current or future programme. They consider the full range of questions and criteria needed to assess a programme's merit and/or worth. Three approaches are categorized as improvement/accountability-oriented studies: Decisions/Accountability – which stresses improvement through serving programme decisions; Consumer orientation – which stresses providing consumers with assessments of optional programs and services; and Accreditation – stressing the need to help consumers to examine the merits of competing institutions and programmes.

Description

The decision/accountability approach denotes evaluation activities that aim to provide programme stakeholders with information required for decision-making relating to initiating, planning, funding, implementing, and using a programme's services on the one hand and information to serve the accountability requirements of the programme. The approach employs the assessed beneficiary needs to collect and characterize the required information. Assessment is done continuously throughout a programme's life-time resulting in dynamic and continuous information flow that may inform the preparation of summative evaluation report once the programme has ended. The evaluator is not required to pass judgment on the programme but to provide a comprehensive facts base that stakeholders may use to make their own decisions. The evaluation involves a wide array of programme stakeholders. The consumer-oriented approach denotes evaluation activities that are focused at judging the relative merits and worth of the products and services of alternative or competing programmes so that consumers have the opportunity to make informed choices about programmes and their services/or products. The approach is built on the premise that consumers, by and large, have no interest in whether the programme goals have been met (or even in what the programme goals are) and are instead more interested in whether their own needs are met. The

accreditation approach denotes studies that aim at assessing institutions, institutional programmes, and personnel to ascertain whether or not they meet requirements of given professions and service areas and are fit to serve designated functions in society. The output of the evaluation is used to serve both accountability and programme improvement requirements.

Method

The accountability/decision-oriented method is contingent on the evaluator maintaining regular interaction with a representative body of stakeholders throughout the evaluation exercise. Through the interaction, the evaluator is able to collaboratively define the evaluation questions, shape the evaluation plans, review draft reports, and disseminate findings. Actual methods that the evaluator may use during the assessment include surveys, needs assessments, case studies, advocate teams, observations, interviews, quasi-experimental / experimental designs; use of checklist; and the application of ethical codes of conduct. Methods of consumer-oriented approach start by compiling a list of topics or areas to be investigated or assessed in the evaluation. The approach recommends that the list of assessment topics must be comprehensive enough to support issuing of decisions on the merit and worth of the programme. Examples of assessment topics for the consumer-oriented methodology include quality; costs; client; consumers; resources; function, delivery system, values, standards, process, outcomes, critical competitors, generalizability, statistical significance and assessed needs. Actual methods that the evaluator may employ include needs assessments, interviews, quasi-experimental / experimental designs, goal-free evaluations, modus operandi analysis, cost analysis, site visits by expert panels, use of checklists; and the application of ethical codes of conduct. The accreditation approach is founded on well established guidelines and criteria that the accreditation body uses to assess merit and worth of the evaluand. The study begins with the evaluand submitting a self report on its performance according to the established guidelines and criteria. The self-report is validated by panels of experts who may also collect additional information through a site visit. In summary, the key methods used in accreditation evaluation are self-reporting/self-study; site visit; and passing of decision by panels of experts.

Evaluation organizer

The decision/accountability approach uses a broad-range of parameters as the evaluation organizers. This includes the persons or body responsible for decision making; the kinds of decision that ought to be made; the accountability requirements placed on the programme; the programme's needs, problems and opportunities; competing approaches; programme operations; programme outcomes; and cost-effectiveness. The consumer-oriented approach likewise employs a broad range of parameters as the evaluation organizers. The parameters closely resemble those of the decision/accountability approach and include programme needs, problems and opportunities;

competing approaches; programme operations; programme outcomes; cost-effectiveness; assessed needs; societal values; and intrinsic criteria of merit. The programme organizers for the accreditation approach are the guidelines and criteria that the accrediting body has adopted; the programme accountability requirements; programme operations; programme outcomes and the intrinsic criteria of merit.

Stakeholder

The decision/accountability approach grants some roles in the evaluation process to a representative body of programme stakeholders that has representation from all stakeholder groups such as programme beneficiaries, parents / guardians, service providers, administrators, consultants, support personnel, policymakers, funding authorities, and citizens. Interaction involves focusing the evaluation and discerning the questions that are of most importance to the stakeholders. No evidence was found to suggest that the consumer-oriented approach involves any stakeholders in the evaluation process. The whole evaluation is performed by the expert evaluator. In the accreditation approach, existing evidence suggests that the individual or institution that is being certified is involved in the evaluation exercise. Typically, the individual or institution is required to complete a self-assessment report, which then constitutes a basis for a detailed assessment by a panel of experts.

Purpose

The main purpose of the decision/accountability approach is to provide a knowledge and value base for making and being accountable for decisions that result in developing, delivering, and making informed use of cost-effective services. This approach strives to continuously provide decision makers with relevant information required to make decisions relevant to the stage of the programme development such as at programme initiation, design, implementation and termination. The main purpose for the consumer-oriented approach is to assist stakeholders make informed choices about products and services. As a result, the consumer-oriented evaluator is urged to focus efforts on judging the relative merits and worth of the products and services of alternative programmes. The main purpose for the accreditation evaluation is to assist certifying bodies to decide whether institutions, institutional programmes, and/or personnel should be approved to deliver specified public services.

Question

The question descriptor provides information on the kinds of issues that are addressed in the evaluation study. In the decision/accountability approach key questions and issues addressed include what consumer needs should a given programme address; what alternative strategies exist for addressing the needs; what plans and resources are required to implement the programme; what is

the health of the programme and whether there is need for a revision; what improvements can be made for better programme operations and results; what is being produced; how does the programme compares with other similar programmes; who is the programme serving and to what effect; is the programme worth the investment. The consumer-oriented approach meanwhile seeks to address several questions including questions around consumer needs and how the needs was being addressed or should be addressed, the operational wellbeing of the programme, how the programme is judged in terms of results and return on investment. For the accreditation approach, the key questions and issues that the evaluation strives to address include what improvement is required and how; what results and outcomes are being produced; and whether the evaluand meets the minimum requirements for accreditation.

This concluding section brings to an end the discussion of the three evaluation approaches characterised under the improvement/accountability category. The discussion is guided by the six descriptors that are developed in section 2.2. On the basis of the description, the following inferences are made:

1. The classification of evaluation approaches as Improvement/Accountability is based on the criterion that the approaches comprehensively appraise a programme for purposes of accountability or to provide information that may be used to improve a current or future programme. On the basis of this criterion, it is observed that the three approaches are a good fit. However, it is noted that the categorization of models as Improvement/Accountability is also driven by an implicit criterion that the model fully assess a programme's merit and/or worth. On the basis of this, the Accreditation/Certification approach appears less likely to provide sufficient information for assessing the merit and/or worth of a programme. Even Stufflebeam notes this weakness when he says that the Accreditation/Certification sometimes puts emphasis on just input and processes. Therefore, only partial perspective of a programme is illuminated which may not be sufficient to measure a programme's merit and/or worth.
2. The three approaches specified under this category all employ multiple organizers to plan and organize the evaluation. The organizers used are either intrinsic to the evaluand (e.g. programme outcome) or external to the evaluand (decision situation or kind of decision that needs to be made).
3. There are approaches under the improvement/accountability category that involve stakeholders in the evaluation while others do not involve any stakeholders. Whenever involved, the role that stakeholders play in the evaluation varies from determining the issues/questions that the evaluation should address to execution of the evaluation and interpretation of its results.

To facilitate further inference, the characteristics of the ten question/method oriented approaches were mapped along the dimensions of whether it uses single or multiple types of organizer; whether the organizers used are intrinsic or external to the programme; the various categories of stakeholders involved in the evaluation; and the role(s) stakeholder(s) play in the evaluation, if any. The outcome of the exercise is summarized in the table below.

Table 4: Characteristic of question/method approaches on organizer & stakeholder descriptors

Approach	Organizer (single / multiple)	Source (Intrinsic / External)	Stakeholder (None / Single / Multiple)	Role (none/ plan/ execute / interpret)
Accreditation	Multiple	External	Single	Execute
Consumer orientation	Multiple	External	None	None
Decisions/Accountability	Multiple	External	Multiple	Plan

There are at least two interesting that can be observed from the information summarized in the above table. The first pattern shows that stakeholder participation in a given evaluation varies depending on whether the organizer used in the evaluation is intrinsic or external to the programme. Where the organizer used is external or both external and intrinsic there is likely to be stakeholder involvement in the evaluation. Although from the above table the Consumer-oriented approach does not involve stakeholders even though it uses organizers that are external. However, this is not surprising given that the consumer-oriented approach is grounded on objectivity and evaluator independence.

2.4.3 Social Agenda/Advocacy Approaches

The social agenda/advocacy approaches seek to ensure that all segments of society have equal access to the opportunities and services provided by a programme. Consequently, the approaches seek to promote democratic principles of equity and fairness by employing programme evaluation to empower those who would otherwise be marginalized. This is achieved by involving all stakeholders in the evaluation and by employing their perspectives in characterizing, investigating, and judging programmes. It is believed that by giving stakeholders the authority to key evaluation decisions such as interpretation and release of findings, evaluators empower them to use evaluation to their best advantage. The four approaches discussed are the client-centred (responsive evaluation); constructivist; deliberative democratic and utilization-focused approaches.

Description

The client-centred/responsive evaluation approach denotes evaluation activities that focus at collecting, processing, and reporting the opinions and judgments of programme stakeholders and those of pertinent experts. For the reason that the perspectives driving the evaluation emanate from various stakeholders; the evaluation does not seek final authoritative conclusion, preferring instead to interpret findings against the divergent value positions held by various programme stakeholders. Doing so is believed to promote equity and fairness and to help people to see things from alternative viewpoints. The constructivists approach denotes evaluation activities that focus on describing knowledge constructions that various stakeholders make on an evaluation question or issue. The approach recognizes that the perspectives and constructions of the various stakeholders are often diverse and sometimes even contradictory; but the evaluator should try to reach consensus with stakeholders on their diverse constructions. The evaluation is conducted by an expert evaluator but with active participation of all stakeholder groups. The deliberative democratic approach denotes evaluation activities that employ democratic participation in the process of arriving at a defensible assessment of a programme's merit and worth. The evaluation is conducted jointly with stakeholders selected on the basis of interest and equity. Stakeholders are charged with the role of providing preliminary findings on key evaluation questions and issues posed. The evaluator, through interaction with stakeholders; reviews all inputs; reflects deeply on all the inputs; then prepares a defensible, well-justified conclusion of the programme. Consequently, although stakeholders are involved, the final conclusion about merit and worth is left to the discretion of the evaluator who is also free to disagree with the views of the stakeholders. The utilization-focused approach denotes evaluation activities conducted on the basis of the intended use to which the evaluation findings shall be put. Although the study is performed by the evaluator, the methods and formats for reporting results are provided for or selected by the intended users. The evaluator conducts the evaluation, including interpretation of findings according to the wishes of the stakeholders or clients. The essence is to ensure that the evaluation addresses questions that are of utmost importance to clients; value positions held by clients are used to interpret findings; preferences held by clients on report formats and content are complied with.

Method

The Client-centred/responsive approach as described by Stufflebeam does not seem to have a definitive method but rather chooses whatever method fits the evaluation purpose. Possible methods include case study; expressive objectives; purposive sampling; observation; adversary reports; storytelling to convey complexity; and socio drama to focus the issues of concern. Data is collected, processed, and reported on a continuous basis to the full range of programme stakeholders. The constructivist approach uses methods such as hermeneutics, dialectical exchange and consensus

development to perform the evaluation exercise. The approach stresses the importance of involving stakeholders and letting them determine the evaluation questions and variables. In the early phase of the study, the evaluator engages each stakeholder group to get their individual constructions on a given evaluation question or issue. This process is likely to result in several individual constructions, which are often divergent. The evaluator then enters into dialogue to discuss with stakeholders the various constructions that have been collected. The essence of the dialogue is to try to reach consensus on the various constructions, although the approach acknowledges that building ultimate consensus is almost an impossible task. For the deliberative democratic approach conducted within an explicit democratic framework of participation, dialogue and deliberation, all stakeholder groups must be equitably represented to participate in the evaluation exercise. The evaluator engages stakeholders and other audiences to compile preliminary findings. The preliminary findings are debated to examine and validate the stakeholders' inputs. In the final stage, the evaluator collates the results from the previous steps into a defensible conclusion about the programme's merit and worth. For the utilization-focused approach Stufflebeam argues that it is open to the use of any credible method of evaluation since it is considered to be situational and dynamic. The evaluator is advised to creatively employ whatever methods are deemed relevant for the problem at hand. Examples of methods that can be used include quantitative and qualitative, formative and summative, naturalistic and experimental. The key tenet of the approach is making client groups responsible for determining the evaluation methods to use. Therefore, the evaluation approach is contingent on what the client group want, and the methods and methodology that they have knowledge of.

Evaluation organizer

The client-centred/responsive approach uses a broad-range of parameters as the evaluation organizers. This includes the concerns and issues that various stakeholders have on the programme; the justification or rationale for the programme; the transactions or operations that the programme is engaged in; the results or outcomes that the programme produces; standards that have been established to guide the programme; and judgement that various stakeholders made on the programme. The constructivist and the deliberative / democratic approaches also employ a broad range of parameters as the evaluation organizers. Parameters used as organizer for the approaches include the concerns and issues that various stakeholders have on the programme; the various perspectives and philosophical standings of the constructivist and deliberative/democratic approaches such as rejection of positivism, democratic participation, and collaborative/unfolding nature of inquiry. The utilization focused approach uses the intended users of the evaluation and the intended use of the evaluation findings as organizers.

Stakeholder

The client-centred/responsive, constructivists and the deliberative democratic approaches all involve the same category of stakeholders in the conduct of the evaluation, although the nature of their involvement varies. The stakeholders who are involved are those responsible for the support, development, administration and operation of the programme under study. For the client-centred / responsive approach, the stakeholders involvement includes providing the issues and questions that the evaluation should answer, validating of the various constructions that have been gathered; and performing the final synthesis of studying and contrasting existing constructions, considering relevant contextual and other information, reasoning out the differences among the constructions, and moving towards reaching consensus. For the deliberative democratic approach, stakeholders are involved in identifying the questions and issues that the evaluation should address and providing preliminary findings on the evaluation questions/issues. The utilization-focused evaluation involves a representative group of stakeholders who are the intended users of the evaluation outcome. Their involvement includes determining the evaluation questions, the methods preferred; the information required by each intended user group; determining formats of the report and helping with interpretation of findings.

Purpose

There are several purposes that the client-centred/responsive approach strives to achieve. These purposes include letting stakeholders know what the entire programme looks and feels; finding out what various stakeholders see as the programme's problems, strengths and weaknesses and how they judge it; and providing stakeholders with the information they need to fulfil their objectives. This approach strives to continuously interact and provide stakeholders with information as they become available. The constructivist approach equally serves similar purposes to the client-centred/responsive approach including continuously searching for key questions and providing stakeholders with useful information as it becomes available; discovering what various stakeholders see as the programme's problems, strengths and weaknesses and how they judge the programme; employing democratic participation in arriving at a defensible assessment of a programme; providing users the information they need to fulfil their objectives. The purposes for the deliberative and utilization approaches are similar. Both approaches strive to employ democratic participation in arriving at a defensible assessment of a programme and to provide users with information they need to fulfil their objectives.

Question

The question descriptor provides information on the kinds of issues that are addressed in the evaluation study. In the client-centred/responsive approach key questions and issues addressed in the evaluation include issues around programme results/impact; programme operation and implementation; stakeholders judgement of the programme; and the justification for the programme. The constructivist and deliberative democratic approaches tend to address similar issues. Issues or questions addressed include whether questions are negotiated with stakeholders and how the various programme stakeholders judge the programme. The utilization-focused approach strives to address questions and issues around programme results/impact; programme operation; and programme costs.

This concluding section brings to an end the discussion of the four evaluation approaches characterised under the social agenda/advocacy category. The categorization of evaluation approaches as Social Agenda/Advocacy category is based on the criteria that the approaches employ the perspectives of all programme stakeholders in characterizing, investigating, and judging programmes. Although participation of programme stakeholder in the evaluation is the key tenet of the criterion; the four approaches adopt varying levels and nature of stakeholder involvement. During evaluation planning; all four approaches use the perspectives of stakeholders in deriving key questions for the evaluation, including bringing to the fore perspectives of the disadvantaged or oppressed groups who are otherwise not listened to. The Utilization-focused approach is however weak on this aspect because its focus is more on use by intended users and less on surfacing the views of the disadvantaged or oppressed groups. However, although the perspectives of clients are taken into consideration in characterizing the evaluation; data collection, analysis and synthesis is the responsibility of the Evaluator. This viewpoint is succinctly captured by Abma & Stake

To be responsive does not automatically yield design authority to stakeholders. It means coming to know the circumstances and problems and values well, then using professional talent and discipline to carry out the inquiry. For me, the inquiry belongs to the evaluator (2001:9).

It is worthwhile to note that although the evaluator is overall responsible for the evaluation, stakeholders still have influence on how data is collected, analysed and reported. In the Client-Centred and Utilization-focused approaches, interpretations and reports are aligned with the values/intended use of/by each stakeholder group. The Deliberative Democratic and Constructivists approaches try to reach consensus with stakeholders on their multiple values and then interprets/reports based on the agreed-on values.

In this section, Stufflebeam's classification of twenty-two evaluation models according to his proposed categories of Question-/Method-Oriented; Accountability-/ Improvement-Oriented and Social Agenda / Advocacy is appraised. Models under each category are assessed against the criteria used to place them into the category. For the question/method-oriented approaches, the criteria is based on whether the models concentrate on answering a set of pointed questions or towards employing specific methods or methodologies; for the Accountability/Improvement-oriented approaches, the criteria is based on whether the models aim to demonstrate accountability or to provide information that may be used to improve a current or future programme; and for the Social Agenda/Advocacy, the criterion is based on whether the approach employs the perspectives of all programme stakeholders in characterizing, investigating, and judging programmes. It is observed that the criteria used for Question-/Method-Oriented and Accountability-/ Improvement-Oriented studies fall short of providing a mutually exclusive way of categorizing evaluation models because models such as those under the Accountability/Improvement-oriented category – could also fit under the Question-/Method-Oriented categorization. After all, every approach to evaluation must have a question and a corresponding method, although it is acknowledged that the questions are at times formulated during the evaluation process. Likewise, many of the models under the question/method-oriented category can fit the Accountability/Improvement categorization. After all, accountability and improvement are two of the three frequently cited purposes for doing evaluation (Rossi, Lipsey & Freeman, 2004; Mark, Henry & Julnes, 2000). There is a suspicion that Stufflebeam was aware of this lack of mutual exclusiveness in his chosen criteria. This was perhaps the reason why there was an implicit criterion of also assessing the model on whether or not it fully assess a programme's merit and/or worth.

The social agenda/advocacy category is noted to also employ double criteria to distinguish its models. First it uses a criterion based on the role that stakeholders play in the evaluation process. A second criterion is based on the premise that evaluation must promote democratic principles of equity and fairness by employing programme evaluation to empower the disenfranchised.

In conclusion, the classification scheme presented by Stufflebeam harbours overlaps and does not provide a mutually exclusive way of grouping models.

2.5 Concluding comments

In this chapter, the classifications of evaluation models performed by Daniel Stufflebeam and Evert Vedung are assessed. Notwithstanding the absence of empirical evidence, there are clear indications that no consensus exists among evaluators on the distinguishing features of an evaluation approach. As Scriven (2003) observed; the urge to denote anything that evaluators do as an evaluation approach appears to be widespread. As a result, every approach presented in the two classifications is not discussed. Approaches such as peer review; pseudo evaluation; objective

testing; outcome monitoring/value-added assessment and performance testing approaches are disregarded from the discussion on the grounds that they do not have their origins in evaluation research or that they are peculiar to a subject-matter area and not the entire evaluation landscape. Although MIS is discussed; there are still question marks over whether it should be considered as evaluation approach. A MIS is an arrangement of people, processes, data and information technology that interact to collect, process, store, and provide as output, the information needed to support an organisation” (Whitten, Bentley & Dittman, 2004, pp.12). This makes MIS more of a supporting tool that is used alongside a credible evaluation approach than being an evaluation approach in itself. For example, the accountability or payment by result approach could benefit from MIS by storing and analyzing the various outcome indicators that the approach collects. In itself, MIS does not evaluate, but can support automation of the particular procedures and data used to conduct monitoring and evaluation.

Related to the foregoing discussion; assessment of the two classifications shows that there appears to be lack of uniformity in naming evaluation approaches. Several scenarios in which approaches share descriptions, but different names are observed. For example Goal-attainment approach (Vedung) vs. Objective-based approach (Stufflebeam); Goal-free (Vedung) vs. Consumer-oriented (Stufflebeam).

Divergent perspectives are also observed in the approaches that seek to measure the achievement of programme goals or objectives. Should the assessment of results be done along the entire continuum of output, outcome and impact? It is concluded that Vedung seems clearer in his description of the goal-attainment and side-effect approaches. These approaches assess results on the entire continuum of output, outcome and impact. It is likely that this must have been the reason why he did not present a separate model for outcome-based assessments. However, this issue is less clear from the descriptions provided by Stufflebeam for the objective-based approach.

CHAPTER 3: ANALYTICAL FRAMEWORK

In the previous chapter, two taxonomies of evaluation approaches are described and discussed. The two taxonomies provide a useful catalogue of existing evaluation approaches and the distinguishing characteristics of each approach. It is observed that some of the approaches are suitable for evaluating programmes or projects while others were less suitable. However, since the focus of this study is on programme monitoring and evaluation information systems, there is a need to identify the distinguishing characteristics/features of evaluation approaches that are suitable for evaluating programmes or projects. This chapter reviews the literature on systems thinking and conceptual modelling. The major outcome of the chapter is an analytical framework developed and used throughout the remainder of the study.

3.1 What is a system?

A system is a *group of interacting, interrelated, or interdependent components that form a complex and unified whole* (Anderson and Johnson, 1997:2). Despite the simplicity in definition, the task of identifying a system and its interrelated parts from a complex problem situation is something challenging. Fortunately, Anderson and Johnson have presented what are considered the five key characteristics of any system. A system; be it physical, biological, designed, abstract, social or organizational must exhibit the following five characteristics (Anderson and Johnson, 1997):

1. All parts of a system must be present for the system to carry out its purpose optimally. As the authors explain, adding or removing parts from a system must have some effects on the system's functioning and/or relationship. If the functioning or relationship of the entity is not destabilized in any way, then the entity should not be considered a system.
2. The parts of a system must be arranged in a specific way for it to carry out its purpose. As the authors explain, arranging the parts of a system in a random order must affect its functioning and/or its relationship. If the functioning or relationship of the entity is not destabilized in any way, then the entity should not be considered a system.
3. Systems have specific purposes within larger systems. As explained by the authors, each system must serve a unique purpose that makes it discrete and unique from any other system. Therefore, it must not be possible to force two or more systems together to get a single, new and larger system.
4. A system must maintain stability through fluctuations and adjustments. As the authors explain, a system is always interacting with its parts and with its environment. The interactions may cause the system to become unstable. A true system must stabilize itself through the interactions, feedback, and adjustments that continually circulate among the system parts, and between the system and its environment.

5. Feedback. A system has feedback within itself and its parts and also within itself and the external systems within which it is embedded. Feedback is not necessarily transmitted and returned through the same system component, and some feedback may return within a short period of time while others may take a long period of time before they return to the system that generated them.

According to Imam, LaGoy and Williams (2006) the systems field has considerably grown from its early past when there were few key concepts and ideas to the current situation where there are thousands of key concepts and ideas. The authors continue to argue that while the expansion in the field enhances its ability to solve complex problems, it also makes it difficult to define precisely and concisely what the field is. Midgley (2006) describes the intellectual development in the systems field using the metaphor of a “wave”. He identifies three “waves” of thinking about systems. In the first “wave” the focus is on improving systems that exist in the real world. To achieve the desired improvement, the physical system is described in a more fundamental way so that “engineers” can design appropriate means of improving it. In the second “wave”, the focus of attention shifts away from improving a physical system that exists in the real world to using systems concepts as tools to a better understanding of the real world. Accordingly the systems that are described in the second “wave” are not regarded as representations of reality but as mental constructions to enable deeper learning about the system. The construction or description of the system is based on views or perspectives solicited from across all the people who have a stake in or are affected by the systems of interest. The third “wave” acknowledges that in society there are always power imbalances and therefore not all perspectives are equal. Consequently each perspective must be subjected to a critique that challenges existing power structures.

There appears to be a pattern of argument that cuts across the discussion in this section. Observation made so far is that all the three “waves” consider the need to describe the system of concern to be a key issue. The points of departure appear to be in the process that is followed to describe the system and the purpose of the description. In the next section additional systems concepts are explored with a view of gaining deeper insight into how systems within a given problem situations are identified described and modelled. The discussion begins with exploration of the various systems thinking concepts that are applicable to programme evaluation.

3.2 Systems thinking and programme evaluation

From the context of an evaluation, systems thinking is an intentional cognitive endeavour, which can only be understood in context of its over-arching purpose. This purpose, according to Branda

is a formal effort to make conscious the systemic construction of a specific concept, situation, problem or evaluand for the purpose of providing the opportunity for the legitimacy of this conceptualization to be questioned, alternative conceptualizations identified, and implications explored (2008:330).

Consequently, systems thinking is not a type of evaluation, but a way of thinking about evaluation, which influences views of what is considered a problem and what its solutions may look like (Branda, 2008; Williams & Iman, 2006).

A central theme in systems thinking is the idea that the entities to be evaluated can be structured or understood as a system (Bawden, 2006; Branda, 2008). Therefore, a systems-based evaluation proceeds by first structuring an evaluand in the form of a system. The task of structuring an evaluand as a system requires evaluators to know systems principles, concepts and theories.

According to Bawden, there are three fundamental principles governing the systems world:

1. *That any whole bounded entity (concrete or abstract, real or assumed) has properties that differ from, and that are unpredictable from a study of its inter-connected component parts.*
2. *That the component parts themselves are systems, and are thus regarded as sub-systems of higher order (supra) systems with their own sub-systems.*
3. *That the emergent properties of all systems are outcomes of the processes of the inter-connections between their sub-systems, and between the systems and the supra-systems in which they are embedded (2006:35).*

There are at least three key points emerging from the fundamental principles listed above. First, both the whole and the parts have properties, but the properties of the whole are not recognizable by studying the parts and their properties. Second, the unique properties of the whole are a result of interactions between the whole and its parts and between the whole and the surrounding systems within which it is embedded. Third, the parts are themselves systems which are referred to as sub-systems of the whole. Therefore, the following keywords run central to the above principle: systems, parts, interactions and properties.

In discussing the above three principles, Bawden (2006) explains that people applying systems approach to evaluation or just any form of inquiry always find themselves working concurrently with

three systems: (a) the system being studied (b) the sub-systems that make up the system being studied, and (c) the supra-system which represents the environment within which the system under study is embedded. From the perspective of programme evaluation, the system being studied is usually considered to be a distinct “programme” or “project”. Taking Bawden’s argument further, a “programme” or “project” is made up of sub-systems which the evaluator must identify during the evaluation. However, being able to identify the system under-study, its sub-systems and its supra-system is just one part of the equation. Other equally important views on systems concepts in evaluation have been discussed in (Imam, 2006; Imam, LaGroy & Williams, 2006; Cabrera, Colosi, and Lobdell, 2008).

Iman (2006) observes that the fields of evaluation and systems thinking are both plagued with communication difficulties and misunderstanding to the extent that no single definition or description of these fields exists. He advises that evaluators wishing to apply systems concepts to evaluation should look to three key patterns, which he believes have the potential to unlock the benefits of systems concepts to evaluation. The three patterns are *perspectives*, *boundaries* and *entangled systems*. He explains that these three patterns can help the evaluator to identify the system of interest, its parts, existing properties (of both the system and the parts) and existing relationships. He explains that *perspectives* requires looking at the object of inquiry from multiple viewpoints and that doing so has potential benefits. His view of *perspectives* is broad and encompasses exploring interconnections between system, its parts and the environment within which the system is embedded. He explains the concept of *boundaries* as an essential element of systems thinking because it provides a way of distinguishing between what is viewed as a system and where its borders lie and what is not. He argues that boundary is fundamentally about values and that its purpose is to scope a given inquiry by clarifying who is inside and who is outside, who will benefit from the study and in what ways. Imam, LaGoy & Willaims (2006) support this viewpoint because to them, setting boundary implies defining what knowledge is considered pertinent, the people who can generate the knowledge, and who has a stake in the results of any attempt to evaluate or to improve the system. For social systems such as programmes, resolving the boundary issue may benefit from the categorization that Jackson (2003) made of the various roles that exist within a purposeful system. The roles within a purposeful system include *decision-maker* – a person who has the powers to make things happen within a system; *actor* – a person who performs a specific task within the system; *customer* – a person who benefits or suffers from what the system does; *problem-owner* – a person who worries about the performance of some aspects of a system; *problem-solver* – a person who tries to improve the system; and *witness* – a person who is affected by the system but is unable to influence it in any way. Regarding *entangled systems*, Imam suggests that systems have very fluid boundaries and that systems are always overlapping or tangling up with one another. In effect, a systems-based investigation must be able to interrogate the situation from various directions and perspectives so that boundary choices are refined and made finer.

Cabrera, Colosi, and Lobdell (2008) advance similar arguments to those discussed above. In the study they develop a pattern of four rules that provide a model of thinking systematically about evaluation: The pattern is abbreviated DSRP and stands for *Distinctions, Systems, Relationships, and Perspectives* respectively. They discuss *distinction* as a process through which a concept is discovered and assigned a distinct identity which makes it unique from other existing concepts. They argue that the identity of a concept is influenced by the context within which the concept is perceived and is deemed meaningful. They reason that concepts only exist in context with other concepts, just as a new idea would likely stem from an existing idea. Therefore, because of the inter-connectedness of concepts, the boundary and identity of a concept only becomes clearer after gradual and continuous reflection on its relationship with the whole and other existing concepts, and after taking on the perspectives of several stakeholders. They state that a concept may also be considered a distinct “whole” that has “parts” or sub-systems. They discuss *systems* as a rule that helps to organize related “parts” or concepts into a “whole”, congruent with the view by Cabrera et al. that “...any collection of related concepts can naturally be viewed as a system” (2008:5). Their view is that the resulting “whole” is also a concept that is amenable to the *distinction* rule discussed above. They also discuss *relationship* as a rule that is employed whenever one makes distinction, formulates systems and attributes perspectives to their systems or concepts. They describe *relationship* as a process of determining how each concept affects the functioning and behaviour of another concept. They also consider *relationship* making as a key process that gives rise to new concepts. Therefore, applying the *relationship* rule implies determining how one concept affects the functioning and behaviour of another concept and vice versa and what new concept emerges from the inquiry. In the DSRP rule, establishing *relationship* between two “parts” or concepts would result into a new and distinct “part” or concept being formed. This new “part” or concept may also be subjected to the *distinction*, *system* and *relationship* rules discussed above. According to the authors, *perspective* is a rule which guides how viewpoints are attributed to an actor. The authors argue that in the DSRP rule, perspectives can be taken from the viewpoint of both human and non-human entities. Therefore, a concept such as “programme activity” can be viewed from the perspective of an implementer; or from the perspective of the environment or setting in which it is being implemented.

This section profiles the perspectives of several theorists on systems concepts and thinking as applied to programme evaluation or to evaluation in general. In the whole, the discussions appear to re-enforce the perspectives of Anderson and Johnson (1997); and Midgley (2006) presented earlier. The preferred approach and way to think about systems appears to be based on pattern rules. This is evident in the discussion by both Imam (2006) and Cabrera, Colosi, and Lobdell (2008). Imam (2006) presents and discusses a pattern rule comprising of three patterns: *perspectives, boundaries* and *entangled system*. Cabrera, Colosi, and Lobdell (2008) also present and discuss a pattern comprising of four patterns: *Distinctions, Systems, Relationship, and Perspective (DSRP)*. The observation at this point is that the pattern rules provided above are all focused at identifying a

system of interest, its sub-system, the parts associated with the system and or its sub-system and the properties that correspond to the system, its sub-systems and its parts. It is also plausible from the discussion that *perspective* and *relationship* just help to re-enforce the discovery and refining of concepts or parts that are applicable to the system of interest.

There is increased interest among evaluators in using systems concepts and approaches in conducting evaluation studies. Imam and Williams (2006) profile several examples of evaluation studies that were conducted using systems-based approaches and concepts. Majority of the studies discuss the application of a particular systems methodology/approach such as soft systems methodology, cybernetic, systems dynamics, dialectic soft systems methodology, critical systems heuristics and human systems dynamics in carrying out an evaluation study. It is important to identify evaluation studies that employed any of the pattern rules discussed above in order to gain better perspective for applying the pattern rules. Luckily, in a study Wasserman (2008; 2010) demonstrates how the DSRP pattern rule can be systematically operationalized for a human service programme.

3.3 DSRP Patterns for a human service programme

Wasserman (2008; 2010) applies the DSRP pattern rule of Cabrera, Colosi, and Lobdell (2008) to evaluate a human service programme. In doing so, she identifies the various systems that make up a human service programme along with their related parts or distinctions. She conceptualizes a human service programme as comprising of two sub-systems: the *provider* and the *target* system; and one supra system; the *human service* system that is formed as a result of the intentional interaction between the *provider* and the *target* systems. For each of the three systems, she identifies the corresponding *distinctions*; *relationships* and *perspectives*. The DSRP patterns corresponding to the *provider*, *target* and *human service* systems are summarised in the table below. The discussion and analysis of the DSRP patterns for the three systems is presented below.

3.3.1 The provider system

The DSRP patter for the *provider* system is described in column two of table 2. As shown in the table, the *distinctions* or “parts” that make up the *provider* system include programme goals and objectives; programme activities; programme providers; administrators; funders; community stakeholders; and macro-environment. According to Wasserman, the above *distinctions* have a hierarchy or order in which they are arranged within the *provider* system. This is very much in sync with the systems concept advanced by Anderson and Johnson (1997). The parts of a system must be arranged in a particular order for the system to function. Wasserman describes the existing hierarchy within the *provider* system as follows:

program activities are nested within the people who provide them (program providers), who are in turn nested in organizations with administrations, all of which are nested within a macro-environment of resources, other programs, policies, practices, norms, etc.) (2008:327).

The hierarchy or order of the *distinctions* within the *provider* appears to be (from highest or most encompassing to lowest or least encompassing): macro-environment -> community stakeholders -> funder -> administrator -> programme provider -> programme activities -> programme goals and objectives.

Table 5: DSRP pattern for a human service programme

System	Provider System	Target System	Human Service System
Distinction between nested parts (from least to most encompassing)	Programme goals and objectives;	Personal goals and objectives;	All distinctions listed in provider and target systems in addition;
	Programme activities;	Existing conditions; Targeted individual(s);	Programme participation;
	Programme providers;	Family, friends and community;	Programme outcomes;
	Administrators;	Macro-environment and relationships between them.	All parts and relationships of the provider and target systems that affect and experience effect of the programme service.
	Funders;		
	Community stakeholders;		
	Macro-environment and relationship between them.		
Operative relationship	Funders to administration;	Targeted individual's effect on existing conditions; effect of existing conditions on targeted individual;	All relationships listed in provider and target systems; In addition:
	Administration to providers;		Participant to programme provider;
	Provider, administrator, and funders to programme objectives;	Target environment affect on existing conditions; effect of existing conditions on target environment.	Participant to programme activities;
	Programme objectives and resources to programme activities;		Participant's environment to programme activities;
			Programme activities to programme outputs and outcomes;
			Provider environment (funders, administration, other

	Providers to programme activity;		stakeholders) to programme evaluation results.
	Funder and administration effect on programme activity;		Participant environment (family, friends, community) to programme evaluation results.
	Programme activity effect on funder, administration and providers.		
Perspectives	View of the various relationships to programme activities and programme objectives	View of the relationships to targeted existing conditions	All operative perspectives listed in provider and target systems
	by:	by:	In addition:
	Programme providers;	Targeted individuals;	View of the value of programme activities;
	Administrators;	Family, friends and community;	View of the value of expected and unexpected outcomes;
	Funders;	Macro-system policy makers and resource providers.	Response to evaluation feedback.
	Macro-system stakeholders		By:
			Programme participants:
			Influential members among programme participants' family, friends or community;
			Programme providers, administrators, funders etc.

Related evaluation research	Organizational development studies;	Needs assessments;	Formative and summative human service programme evaluations.
	Performance evaluations;	Risk and asset assessments;	
	Programme monitoring.	Behavioural research;	
		Epidemiological studies.	

Source: Wasserman, 2010

Wasserman also identifies several operative relationships for the provider system. These relationships include funders to administration; administration to providers; provider, administrator, and funders to programme objectives; programme objectives and resources to programme activities; providers to programme activity; funder and administration effect on programme activity; programme activity effect on funder, administration and providers. During an evaluation, these relationships are interrogated from the perspectives of various stakeholders. However it may not be necessary to interrogate all the above-listed relationships during a particular evaluation study. Wasserman (2008) advises that for evaluation of the *provider* system; the relationships that need to be interrogated are the relationships to programme activities and programme objectives. She advises that the relationships should be interrogated from the perspectives of programme providers; programme administrators; programme funders; and macro-system stakeholders. She reasons that evaluation of the *provider* system falls within the realm of organizational development, performance evaluation and programme monitoring studies.

3.3.2 The target system

The DSRP patter for the *target* system is described in column three of table 2. As shown in the table, the *distinctions* or “parts” that make up the *target* system include personal goals and objectives; existing conditions; targeted individual(s); family, friends and community; macro-environment and relationships between them. Just like the *provider* system, the *target* system also has its distinctions in a hierarchical manner. To quote Wasserman:

In the target system, certain conditions exist within or around a targeted individual who is nested within a family and community, which in turn is also situated in a macro environment of practices, programs, policies, resources and norms, some overlapping but often independent of those affecting the program (2008:327).

The hierarchy or order of the *distinctions* within the *target* system appears to be (from least or less encompassing to highest or most encompassing): existing conditions -> targeted individual -> family and community -> macro environment. Operative relationships that Wasserman identifies for the *target* system include targeted individual's effect on existing conditions; effect of existing conditions on targeted individual; target environment affect on existing conditions; effect of existing conditions on target environment. For evaluation of the *target* system Wasserman (2008) advises that evaluators should focus at interrogating the various relationships to targeted existing conditions from the perspectives of targeted individuals; family, friends and community; macro-system policy makers and resource providers. She reasons that evaluation of the *target* system falls within the realm of needs assessment; risk and asset assessments; behavioural research and epidemiological studies.

3.3.3 The human service system

The DSRP pattern for the *human service* system is described in column three of table 2. As shown in the table, the distinctions or “parts” that make up the human service system include the distinctions listed in provider and target systems; programme participation; programme outcomes; parts and relationships of the provider and target systems that affect and experience effect of the programme service. The hierarchy or order of the distinctions within the human service system, although not explicitly indicated should inherit from the two sub-systems. Exception being programme participation and programme outcomes that come into being as a result of the provider system’s intentional relationship with the target system. Operative relationships identified for the human service system include relationships listed in the *provider* and *target* systems; participant to programme provider; participant to programme activities; participant’s environment to programme activities; programme activities to programme outputs and outcomes; provider environment (funders, administration, other stakeholders) to programme evaluation results; participant environment (family, friends, community) to programme evaluation results. For evaluation of the *human service* system Wasserman (2008) advises that evaluators should focus at interrogating the various relationships as specified in the *provider* and *target* systems in addition to view of the value of programme activities; view of the value of expected and unexpected outcomes; response to evaluation feedback from the perspectives of programme participants; influential members among programme participants’ family, friends or community; programme providers, administrators and funders. She reasons that evaluation of the *human service* system falls within the realm of formative and summative human service programme evaluation.

This section describes the DSRP patterns that Wasserman developed for a human service programme. The description covers the three systems that make up a human service programme: the provider, target and human service systems. It is observed that each of the system may be studied independently. Independent studies of the provider systems falls within the realm of organizational development, performance evaluation and programme monitoring studies. Likewise independent study of the *target* system falls within the realm of needs assessment; risk and asset assessments; behavioural research and epidemiological studies. Any study of the *human service* system falls within the realm of formative and summative human service programme evaluation. The study of each system involves interrogating the various relationships that exist between/among the distinctions of the system. The relationships are viewed or studied from the perspectives of those who have interest in or are affected by the system of concern. For the provider system, relevant stakeholders whose perspectives are sought include programme provider; programme administrator; programme funder; and macro systems stakeholders. For the target system relevant stakeholders include targeted individuals; family, friends and community; macro-system policy makers and resource providers. For the human service system, the relevant stakeholders whose perspectives are sought include

programme participants; influential members among programme participants' family friends or community. Other than the human service system that is born from the interaction of the both the provider and the target system; the provider and the target systems individually have unique distinctions, relationships and perspectives.

3.4 Conceptual modeling theory

Information systems collect, store, maintain, and disseminate knowledge about physical and social worlds. Our context is social world of programme monitoring and evaluation; and the study aims to develop an accurate representation of this world. A starting point in creating a conceptual model for programme monitoring and evaluation information system is to identify the set of constructs that exist in the world of M&E. Section 3.2 discusses various pattern rules that are useful in conceptualizing systems. However, pattern rules are not the only theories that guide in conceptualizing things. Research has demonstrated that theory of ontology is useful in identifying the important constructs in any real-world domain (Weber, 2003; Wyssusek, 2006).

A theory of ontology postulates that the world is made up of *things* that possess *properties* Wand *et al.* (1995). A property can be *intrinsic* to a thing or *mutual* to several things. For example, colour is an *intrinsic* property of a person and it exists irrespective of whether or not its existence is recognized. However, marital status is mutual to a person and marriage. In conceptual modelling, the terminology *attribute* is used to denote characteristics that humans assign to things. An attribute may correspond to an existing property of a thing, but not every attribute has to represent a property. In this thesis, it is preferred to use the terminology attribute instead of *property* because it encompasses all kinds of properties, including the ones that are *mutual*.

3.5 The Analytical Framework

Earlier in this chapter, Anderson and Johnson (1997) present a succinct discussion of what a system is and what its essential characteristics are and show how the idea of "thinking" about system influences methodological and theoretical development in the evaluation field. A common approach to applying systems thinking to evaluation is the idea of structuring the entity to be evaluated into a system. This structuring implies conceptualizing the entity into a whole and parts that are logically linked and inter-related. Key strategies for conceptualizing things into whole and inter-related parts include the use of pattern rules (Iman, 2006; Cabrera, Colosi, and Lobdell, 2008) and theory of ontology (Weber, 2003; Wyssusek, 2006). Therefore, evaluation within this context involves:

1. Identifying and describing all the "parts" that makes up the programme of interest. A programme is made up of several "parts". Section 3.2 describes how pattern rules could be

used to identify and describe the essential “parts” of a programme. Likewise section 3.4 describes how theory of ontology could be used to identify and describe the essential “parts” of a programme. Examples of the essential “parts” that correspond to a human service programme are described in section 3.3. In the description, the term *distinction* is used instead of “parts”. However, the *attributes* associated with the *distinction* described in section 3.3 are not specified. And yet according to a theory of ontology a complete description of a “part” should include a description of its *attributes*. Therefore, a conceptual model for a programme monitoring and evaluation information system must be designed based on the various distinctions that comprise a programme and its evaluation and the associated attributes. In this regard “distinction” and “attribute” should be considered in the analytical framework.

2. Arranging the parts in a way that recognizes the hierarchical relationships among the parts, including identifying the sub-systems there-in. This process helps to clarify the dependencies that exist between various parts. In section 3.3, some rudimentary specification of existing dependencies between the parts in *provider*, *target* and *human service* systems is presented. For *provider* system, the dependency is provided as (from highest or most encompassing to lowest or least encompassing): macro-environment -> community stakeholders -> funder -> administrator -> programme provider -> programme activities -> programme goals and objectives. In this example, the dependency between programme provider and programme activities can be described as “there will be no programme activity if there is no programme provider”. It should be noted that there could be several ways that two “parts” depend on one another. For example “there shall be several programme activities for a single programme provider”. Therefore, an important consideration for the analytical framework is “relation” because it informs the way concepts should be arranged within a conceptual model for programme monitoring and evaluation information system.
3. Identifying relationships between parts and interrogating them from the perspectives of people whose views are considered important in gaining better insight into the functioning of the relationship and its effect. As discussed in section 3.2 each time this process is undertaken, new “parts” may emerge or existing parts are refined. Section 3.3 lists several operative relationships and the perspective from which they could be interrogated. However, there is need to make consideration for this in the analytical framework since the outcome of this process that is considered important for developing conceptual model for programme monitoring and evaluation information system shall emerge from (1) and (2).

From the foregoing discussions, it is possible to expound the essential elements of the analytical framework proposed in this chapter. The analytical framework is based on the discussions in sections 3.1; 3.2; 3.3 and 3.4. The framework comprises three major dimensions:

- **Distinction:** this dimension of the framework describes a key construct that is identified as being a constituent part of the programme under investigation. Examples of *distinction* include programme activity; programme provider; and programme administrator. In section 3.3; the complete list of *distinctions* applicable to human service programme is provided. While the list may present an incomplete picture for other types of programmes, it nevertheless provides a starting point for the study.
- **Attribute:** this dimension of the framework describes the properties of each identified *distinction*. A property is a unique characteristic of a “thing” and collectively they are what make one *distinction* different from another *distinction*. There are usually several attributes for a single *distinction*. As relationships between different *distinctions* are interrogated from various perspectives, the *distinctions* that are involved in the relationship may acquire new attributes that shall also need to be described.
- **Relation:** this dimension of the framework describes the hierarchical nature of relationships that exist between two or more *distinctions*. As discussed earlier, *distinctions* are of varying stature, which means some can be viewed as “slaves” and others as “master”. In the whole, *relation* is the glue that keeps the parts of the system together.

In this chapter, literature from two distinct, though related fields: systems thinking and conceptual modelling is presented and discussed. Congruent with the aim of this research: to develop a conceptual model for a monitoring and evaluation information system; reviewing literature on systems thinking is critical as it enables the demarcation of the important patterns considered to be generic for programme monitoring and evaluation information system. These constructs are found to include programme goals and objectives; programme activities; programme providers; administrators; funders; community stakeholders; macro-environment and relationship between them; personal goals and objectives; existing conditions; targeted individual (s); family friends, and community; macro-environment and relationships between them; programme participation and programme outcomes. Through a further discussion of literature on conceptual theory, it is noted that conceptual models are constructed from key constructs, their attributes and relations. Using this information base, an analytical framework is developed to guide the collection and analysis of data. This framework is employed in chapter 5 to identify and collect the necessary data required for the development of the conceptual model.

CHAPTER 4: DESIGN AND METHODOLOGY

The previous two chapters provide some general background to the main objectives of this study. In chapter 2; the key M&E concepts that characterize various evaluation approaches are identified and discussed. In chapter 3; the programme-specific M&E concepts are isolated from the general list developed in chapter 2. To a great extent, this information base provides a framework to guide the empirical component of this thesis.

In this chapter, the procedures and guidelines used to collect and analyse data for addressing the research questions are presented. The chapter also describes the unit of analysis; sampling methods; and in general the limitations of the study.

In chapter 1, the key question of the research is formulated in the following way: *how can software for monitoring and evaluation be designed to allow adaptability across different programmes?* Following the discussion of the literature in chapters 2 and 3; it is evident that the design of adaptable M&E software is influenced by one's understanding of various M&E concepts such as outcome; programme goal; and programme objective - and the interactions between and among them. In both the practice and theory of M&E; these issues tend to be operationalized differently across programmes. Since a key focus of this thesis is to merge these divergent operationalizations of the key M&E concepts into a conceptual model; the initial research question is revised to reflect this perspective. The guiding research question evolves to become:

How can the key concepts used in programme monitoring and evaluation be arranged to enable the development of adaptable programme monitoring and evaluation software?

4.1 Unit of Analysis

This thesis aims to develop a conceptual model for a programme monitoring and evaluation information system. Such a model, it has been argued, should contain the essential information required to design and develop adaptable software for programme M&E. "Programme" is chosen as the unit of analysis. Therefore, the data collected during the thesis is derived from M&E documents of selected case-study programmes. Each case-study programme is described in terms of its intention, the targeted beneficiary, time period, status at the time of the study, implementing organization and information on delivery mechanism or strategy adopted.

4.2 The Analytical framework

This section introduces the analytical framework developed in chapter 3. The analytical framework describes the essential category of information that is required for developing a conceptual model for a programme monitoring and evaluation information system. From literature, it is discovered that a common approach to applying systems thinking to evaluation is the idea of structuring the entity to be evaluated into a system. The resulting “structure” would comprise of: (1) the various parts that constitute the programme; (2) the properties that are unique to each part; and (3) the order or hierarchy of relation between the various parts. In the thesis, it is preferred to name the various parts that constitute the programme’s essential elements “Distinction” and the properties unique to each part “Attribute”. Also, it is preferred to name the inter-dependency of parts “Relation”. Consequently, the essential elements of the analytical framework include “Distinction”, “Attribute” and “Relation”.

4.2.1 Distinction

“Distinction” describes any key construct that is identified as being a constituent part of the programme under investigation. It is an important element because it determines the focus of the evaluation. Evaluation of programme typically involves querying relationships that exist between two or more distinctions. In the study, all the distinctions belonging to each case study programme shall be identified and described. The distinctions identified from each case study programmes shall then be used to develop the conceptual model proposed in this study. In chapter 3 a list of distinctions which are generic to a human service programme is presented. The list includes: programme goals and objectives; programme activities; programme providers; administrators; funders; community stakeholders; macro-environment and relationship between them; personal goals and objectives; existing conditions; targeted individual (s); family friends, and community; macro-environment and relationships between them; programme participation and programme outcomes. This list shall be used to guide the identification of distinctions in each case study programme.

4.2.2 Attribute

The “Attribute” dimension describes the properties or characteristic of a *distinction*. A property is a unique characteristic of a “thing” and collectively they are what make one *distinction* different from another *distinction*. There are usually several attributes for a single *distinction*. An “Attribute” can be intrinsic or mutual. An intrinsic *attribute* exists irrespective of whether or not its existence is recognized. For example, colour is an intrinsic property of a person. However, marital status is mutual to a person and marriage. Therefore, unless a “Male person” relates with a “Female person” there won’t be any property of a person called marital status. This example also explains the notion that

attributes may form as a result of the interaction between two or more distinctions. In the study, both the intrinsic and mutual *attributes* of each *distinction* were identified. The mutual property of *distinctions* is identified by interrogating relationship between and among several *distinctions*.

4.2.3 Relation

“Relation” describes the nature of the dependency between *distinctions*. *Distinctions* have varying statue, which means some *distinctions* could be viewed as “slaves” while others are viewed as “master”. In the whole, *relation* is the glue that keeps the parts of the system together. For example, given that programme provider and programme activities are two *distinctions* that depend on one another, and that the dependency is such that a programme activity is lower in statue than programme provider; the dependency could be described as “there will be no programme activity if there is no programme provider”. It should also be noted that there could be several ways that any two distinctions depend on one another. Again, using the earlier example of programme provider and programme activity; another dependency could be described as “there shall be several programme activities for a single programme provider”. Therefore, the way that identified *distinctions* relate with another shall be described for each case-study programme.

4.3 Sampling

The sample used in this study comprises of 15 monitoring and evaluation study reports of three large-scale programmes implemented in Uganda. The documents are drawn from the development experience clearinghouse online database (<http://www.dec.usaid.gov>). This database contains over 16,000 electronic documents of various studies conducted or commissioned by United States Agency for International Development (USAID) around the world. Selection of the documents is guided by purposeful sampling techniques using the following criteria:

- The selected document corresponds to a programme that was implemented in Uganda. This criteria ensures that work is within a familiar context, which makes analysis much faster and accurate since the core issues discussed in the documents could be easily understood and apprehended;
- The selected documents are for health-related programmes. This criterion is influenced by the current world-wide attention accorded to the health sector. Focusing on the health sector ensured that work is within a sector that had relatively advanced M&E practices;
- Selected documents belonged to one of the following types: programme plan; progress report, mid-term evaluation; final evaluation; and periodic M&E studies. The rationale for this

criterion is the belief that documents of these types address issues of monitoring and evaluation; and therefore will contain the key concepts being sought in the study;

- Selected documents are less than ten years old. M&E is a fast and dynamic discipline with new ideas emerging quiet frequently. Therefore, basing the study on recent documents ensures that current thinking in the field was captured.

During sampling, all documents satisfying the above criteria are grouped according to their corresponding programme. Three case study programmes are selected from the list on a purposeful technique. Each selected programme also has to have at least four documents grouped under it. The three selected programmes are: The AIDS/HIV Integrated Model district programme (AIM); The Northern Uganda Malaria Aids Tuberculosis programme (NUMAT) and the President's Emergency Plan for Aids Relief (PEPFAR). A summary of the programmes and the associated documents is provided in the table below.

Table 6: Case study programme and the documents sampled

Programme	Document	Document Type	Year
AIM	AIM Project evaluation	End-of-Programme Evaluation	2007
	End of Programme Report	End-of-Programme Report	2006
	Annual Report for AIM Programme Year 1: July 2003 – September 2004	Annual Progress report	2004
	Annual Report for AIM Programme Year 2: July 2002 – June 2003	Annual Progress report	2003
	Annual Report for AIM Programme Year 1: June 2001 – July 2002	Annual Progress Report	2002
NUMAT	Mid-term review of the NUMAT Programme	Annual Progress Report	2009
	Annual Report for NUMAT Programme Year 2: October 2007 – September 2008	Annual Progress Report	2008
	Annual Report for NUMAT Programme Year 1: October 2006 – September 2007	Mid-term evaluation	2007

Programme	Document	Document Type	Year
PEPFAR	Next generation Indicators Reference guide	Reference Guideline	2009
	PEPFAR Public Health Evaluation – Care and Support	Periodic Evaluation	2009
	Country Operating Plan 2009: Uganda	Annual Programme Operational plan	2009
	Country Operational Plan (COP) Guidance: Programmatic Considerations	Planning & reporting Guideline	2009
	Country Operating Plan 2008: Uganda	Annual Programme Operational plan	2008
	Country Operating Plan 2007: Uganda	Annual Programme Operational plan	2007
	Audit of USAID/Uganda's implementation of PEPFAR	Periodic evaluation	2005

4.4 Data collection and analysis

The major sources of data for the study are described in section 4.3 and include end-of-programme evaluation reports; mid-term evaluation reports, annual progress reports, reference guidelines, periodic evaluation reports, planning/reporting guidelines and annual programme operational plans. These sources represent qualitative data, and are amenable to qualitative methods of data collection and analysis. In this thesis, data collection and analysis was performed concurrently and in a recursive and iterative manner.

The process begun with repeated reading of the various documents associated with the selected case study programmes. This initial reading was intended to familiarize and to gain insight into the data. No particular order was followed and documents were read on an ad hoc basis. The end result

of this initial process was a broader understanding of the common issues in the documents along with the relevant contextual information associated with each case study programme. Afterwards, identification and extraction of text considered relevant to the study was performed.

The identification and extraction of the relevant text was guided by the analytical framework proposed in chapter 3. Specifically, the three dimensions of the framework: distinction, attribute and relation were utilized. For each dimension, relevant texts were extracted from the source documents and described.

Analysis involved determining the relationships that existed between and among different distinctions. Relationships were assessed in the form of the constraints that one distinction placed on the existence of another distinction. Next, conceptual models were developed to represent the order and relationship of distinctions for each case programme. These individual conceptual models were further analysed in a comparative manner with the result being a generic conceptual model that combined features of the individual models.

4.5 Limitation of the study

The conceptual model presented in this study was developed from information gleaned from evaluation-related documents, including evaluation reports, progress reports, and end of programme reports without the participation of the would-be model users and domain experts. The literature on conceptual modelling recommends that domain experts and the would-be model users should be involved in the development process. Such broad-based and inclusive process ensures that a common perception of reality is attained among observers of the system. Due to time factors, neither the would-be users nor the domain experts were consulted. Such consultations would have helped to validate the model and to further refine it. As a result, the model developed in this study might have missed certain essential elements relating to the world of monitoring and evaluation information system.

CHAPTER 5: RESULTS

This chapter reviews the practice of monitoring and evaluation within three large scale programmes implemented in Uganda. In particular the review analyzes programme plans, progress reports, evaluation reports and planning/reporting guidelines of the three case study programmes. The essence of the analysis is to identify the elements/components of a programme that are essential for planning, monitoring, evaluation and reporting. The elements/components that emerge from the analysis provide input into development of the conceptual model for programme monitoring and evaluation information system performed in the study.

The three case study programmes selected for the study include (1) the AIDS/HIV Integrated Model District Programme (AIM); (2) the Northern Uganda Malaria AIDS Tuberculosis Programme (NUMAT); and (3) the President's Emergency Plan for AIDS Relief (PEPFAR). These three programmes were selected on the basis of a purposeful sampling technique. Several documents were drawn from the development experience clearinghouse online database (<http://www.dec.usaid.gov>) based on a set of criteria. The guiding criteria was that the document matches a health-related programme implemented in Uganda; describes a programme plan / progress report / mid-term evaluation / final evaluation / periodic M&E studies; and was a recent publication of not more than ten years old. The three programmes that got the highest number of matched documents were selected for the study.

5.1 The AIDS/HIV Integrated Model District Programme

The AIDS Integrated Model District (AIM) Project was a multi-faceted effort implemented in 16 districts of Uganda from 2001 to 2006. It was designed to assist with increasing the provision of quality integrated HIV/AIDS services at districts and sub-districts so that men, women and children in the selected districts could access and utilize appropriate, affordable and quality integrated HIV/AIDS/TB prevention, care, and support services. To achieve this, AIM provided support to national and local level NGOs active in the HIV/AIDS prevention, care and support efforts, and to Ugandan government ministries and agencies at the national and district levels. AIM was delivered mainly through a system of grants to local- and national-level civil society organizations and government entities.

5.1.1 Description of distinction

In chapter 3, insight from literature was utilized to identify and discuss the key M&E constructs that are applicable to programme monitoring and evaluation. In this section, the particular M&E constructs that were utilized in the AIM programme are identified and described. Each identified construct is described in terms of its key attributes.

Programme Goal

Within the various M&E documents associated with AIM, one programme goal was identified: for men, women and children in sixteen selected districts in Uganda, to access and utilize appropriate, affordable and quality integrated HIV/AIDS prevention, care and support services. Through a critical analysis of this goal statement, the following attributes were derived:

- Goal-statement: this attribute represents the text assigned to the programme goal. For example; men, women and children in sixteen selected districts in Uganda, to access and utilize appropriate, affordable and quality integrated HIV/AIDS prevention, care and support services;
- What-to-improve: this attribute describes the situation that the programme intends to make better by its various interventions (shown in bold in the two examples). For example; **access to** appropriate, affordable and quality integrated HIV/AIDS prevention, care and support services; **utilization of** appropriate, affordable and quality integrated HIV/AIDS prevention, care and support services.
- Who-benefits: this attribute describes the person or entity that is expected to benefit from the improvement being sought. For example; men, women and children in sixteen selected districts in Uganda;

Programme Objective

Within the various programme documents, five objectives of AIM were identified: (1) To strengthen and support the capacity of government, non-governmental organizations (NGOs), community based organizations (CBOs), faith-based organizations (FBOs) and the private sector to plan, implement, manage and provide quality services at the national, district and sub-district level; (2) To increase integration and quality of comprehensive HIV/AIDS prevention, care and support services in selected districts; (3) to increase access to and utilization of quality HIV prevention services in selected districts and sub districts; (4) to increase access to and utilization of quality HIV/AIDS clinical, community and home-based care in selected districts and sub-districts; and (5) to increase access to

and utilization of quality social support services for people infected and affected by HIV/AIDS including orphans, vulnerable children and adolescents in selected districts and sub districts.

Through a critical assessment of these five programme objectives, the following attributes were derived:

- Objective-statement: this attribute represents the text assigned to the programme objective. For example; to increase access to and utilization of quality HIV prevention services in selected districts and sub districts;
- What-to-improve: this attribute describes the situation that the programme intends to make better by implementing the stated objective. For example; capacity to plan, implement, manage and provide quality services; integration and quality of comprehensive HIV/AIDS prevention, care and support services;
- Who-benefits: this attribute describes the person or entity that is expected to benefit from the improvement being sought. For example; districts and sub-districts; people infected and affected by HIV/AIDS including orphans, vulnerable children and adolescents;

Programme Activity

The description of PROGRAMME ACTIVITY was contained within the various annual reports. The annual reports documented the activities that were carried out during every reporting period. The activities represent AIM's concerted attempts to ameliorate the problem situation that it was expected to affect. Activity reporting was a mean by which AIM showed progress towards attainment of established goal and objectives. Through a critical assessment of the activities described within the annual reports, the following attributes were derived:

- Activity-Category: this attribute describes the broad area of intervention or service to which a given activity belongs. Example of activity category include HIV/AIDS needs assessments and strategic plans; Referral network for AIDS (RNA).
- What-was-done: this attribute describes the actions that were performed during the reporting period. The description is in narrative form and may sometimes mention the approach used in conducting the activity and the outputs produced. An excerpt from the AIM annual reports is shown below:

AIM contracted a local firm to conduct a needs assessment exercise in the Phase I districts. Though the exercise took longer to complete than anticipated, nine district reports had been completed by the end of 2002. Copies of these reports are available in the AIM resource centre

- Output: this attribute describes the product or service that has resulted from the activity that AIM carried out during a given reporting period. In describing the output, a quantitative measure (usually numerical) is provided to indicate the quantity of the output that has been produced. Excerpt from the AIM annual reports is provided below:

nine district reports had been completed by the end of 2002. Copies of these reports are available in the AIM resource centre

- Adjustment: this attribute describes any changes that have been made to an activity, and may sometimes also incorporate a description of the reasons for the changes. An excerpt from the AIM annual reports is shown below:

... it would be logistically difficult for AIM to monitor their activities if they were funded by the programme given the difficulty of travelling to the district. It was therefore resolved to go into partnership with some national level organizations that had experience of working in conflict situations and that were in position to partner with local organizations in Pader

- Reason-for-Adjustment: an attribute that describes the rationale for to the changes that have been made to an activity, if any. An excerpt from the AIM annual reports is provided below:

During the year, it became apparent that there weren't a significant number of organizations that could provide HIV/AIDS services in Pader district, and even for the few that were there it would be logistically difficult for AIM to monitor their activities if they were funded by the programme given the difficulty of travelling to the district. It was therefore resolved to go into partnership ...

- Challenge: an attribute that describes the problems that were encountered during activity implementation. An excerpt from the 3rd annual report is shown below:

Key challenges to VCT implementation include inadequate staff, irregular supply of HIV test kits and insufficient systems to follow up people who have been tested.

- Remedy-to-the-challenge: an attribute that describes the proposed solutions to the challenge described above. An excerpt from the 3rd annual report is shown below:

key challenges to VCT implementation include inadequate staff, irregular supply of HIV test kits and insufficient systems to follow up people who have been tested.....AIM has continued to work with MOH, DELIVER and NMS to improve logistics management through training of staff and regular technical assistance;

- Lesson-Learnt: an attribute that describes any new knowledge acquired as a result of the activity.
- Time-Period: this attribute specifies the frequency with which AIM was required to submit activity implementation report. For AIM, activity reports were submitted on an annual basis. But reporting frequency may be set to any time period including weekly; monthly; and quarterly.

Programme Provider

The primary means through which AIM delivered services was grants to national and local level NGOs active in the HIV/AIDS prevention, care and support efforts, and to Ugandan government ministries and agencies at the national and district levels. The following were the key types of programme providers: civil society organisations (CBOs; FBOs; NGOs; PSOs) and Ugandan government entities (DDHS; DAC; HSD). This is captured succinctly by the AIM final evaluation:

AIM was designed to work closely with local government, NGOs, CBOs, FBOs, the private sector and other partners to increase availability and access to a range of core comprehensive HIV/AIDS services in selected districts.

Through a critical assessment, the following attributes were derived:

- Provider-name: an attribute that may represent a unique identifying characteristic of a programme provider.
- Provider-type: an attribute that may represent the kind of organization that each provider belongs. Recognized types include CBO; FBO; NGO; PSO; and Government;

Programme Administrator

AIM was implemented by JSI Research Training Institute, Inc., and included World Education and World Learning as consortium members. However, the actual programme administration was structured around three entities: the District HIV/AIDS Committee (DAC), the AIM's regional offices and the AIM's national office. The DACs were responsible for HIV/AIDS activities that occurred at the district level, including supporting the soliciting, reviewing, selecting and awarding of grants. At regional level, the AIM's regional office provided oversight to a cluster of districts within a single geographical region. The overall oversight for the programme rested with the AIM's national office located in Kampala. Through a critical assessment, the following attributes were derived:

- Implementing-Agency: an attribute that specifies the name of the main organization or entity responsible for implementing the AIM programme. E.g. JSI Research Training Institute.
- Implementing-Partner: an attribute that specifies the names of all other organizations or entities supporting the implementing agency. E.g. District HIV/AIDS Committee (DAC).

Programme Funder

The initial funding for AIM came from USAID and CDC. Later, AIM became a prime partner in the emergency plan for aids relief (PEPFAR) and subsequently begun to receive PEPFAR funds. An excerpt from the final evaluation of AIM is shown below.

Initially, funding for this programme came through USAID and CDC. In November 2002, the ceiling for the cooperative agreement was increased from \$19 million to \$38 million through funding from the LIFE Initiative and USAID infectious disease resources for tuberculosis. CDC funding ended in programme year three, and since January 2004, AIM has been funded directly by USAID through the President's Emergency Plan for AIDS Relief (PEPFAR) initiative.

The following attributes have been derived for this distinction:

- Funder-name: an attribute that may represent a unique identifying characteristic of the entity financing the programme. For example PEPFAR; USAID; CDC
- Funding-amount: an attribute that may represent the sum of money that the Funder is committing or making available for the programme. For example; \$12m.
- Funding-period: an attribute that may represent the time period within which the fund is made available; e.g. 2002-2007.

Target Population

AIM targeted several categories of beneficiaries. There are three main types of beneficiaries that AIM supported:

- District Aids Committees (DACs): according to the final aim evaluation report, AIM planned to strengthen the capacity of DACs to be effective players in the management of decentralized health service delivery by building skills of committee members in problem analysis, prioritization, planning, coordination and accountability.
- Youth and Adult: AIM planned to reach the youths and adults in the programme districts with HIV/AIDS prevention and care/support services.

- HIV+ pregnant women: this population group was targeted mainly with the prevention of mother to child transmission programming.

The following attributes have been derived for this distinction:

- Target-population: an attribute that represents a unique identifying characteristic of the people or entity targeted by the programme. For example Youth; Adult; HIV pregnant women.
- Service-provided: an attribute that specifies the products that the programme provides to each target population. For example Prevention of Mother to Child Transmission.

Programme Outcome

AIM expects to affect its target population in varied ways and 11 outcomes associated with AIM were identified. They include: (1) Increased access to and utilization of strategic information; (2) Increased access to and utilization of interventions for OVC; (3) Increased access to and utilization of quality HIV/AIDS clinical, community and home-based care; (4) Increased access to and utilization of quality TB/HIV prevention and treatment services; (5) Improved laboratory capacity for HIV/AIDS/TB; (6) Increased access to and utilization of prevention services for at-risk populations; (7) Increased access to and utilization of quality STI services; (8) Increased access to and utilization of quality PMTCT services; (9) Increased access to and utilization of quality HCT services; (10) Established/strengthened capacity to deliver comprehensive and integrated HIV/AIDS interventions among district, government, NGO, CBO, FBO and Private Sector; (11) Established/strengthened district HIV/AIDS planning, monitoring and feedback processes.

Assessment of the eleven outcomes reveals that the following attributes are associated with programme outcome:

- Outcome-statement: this attribute represents a description of a particular outcome. For example; Increased access to and utilization of strategic information;
- Anticipated-change: this attribute represents the nature of improvement that has resulted or is expected to be achieved. Increased access to and utilization of strategic information. However, it was noted that the outcome statement and the anticipated change were in most cases identical;
- Who-changes: this attributes describes the person or entity that experienced or is expected to experience the anticipated change. For example; at-risk populations (reference outcome number 6).

Programme Management Plan

The Programme Management Plan (PMP) is USAID's principal tool for managers to monitor progress of their programmes and to make the necessary adjustments. The PROGRAMME MANAGEMENT PLAN combines attributes from PROGRAMME GOAL; PROGRAMME OBJECTIVE and PROGRAMME OUTCOME. The following are its attributes:

- Goal-statement: this attribute was derived from the PROGRAMME GOAL distinction. It relates to a goal that had already been defined there. The essence is to tie specific elements of the PMP to a specified programme goal.
- Objective-statement: this attribute was derived from the PROGRAMME OBJECTIVE distinction. It relates to an objective that had already been defined there. For example; to increase access to and utilization of quality HIV prevention services in selected districts and sub districts. The essence is to tie specific elements of the PMP to a specified programme objective.
- Outcome-statement: this attribute was derived from the PROGRAMME OUTCOME distinction. It relates to an outcome that had already been defined there. The essence is to tie specific elements of the PMP to the expected and/or realized outcome.
- Performance-measure: this attribute defines the measure that is used to sense achievement of the outcome specified in the outcome-statement (see above). For example; Number of AIM-supported districts that track HIV/AIDS indicators.
- Baseline-value: this attribute specifies the value of the performance measure prior to AIM.
- Target-value: this attribute specifies the expected increase or decrease in the performance measure and by what magnitude.
- Actual-value: this attribute specifies the observed increase or decrease in the performance measure and the corresponding magnitude.
- Time-Period: this attribute specifies the frequency with which progress on the implementation of the PMP is provided. Under AIM, it was provided on an annual basis. But other frequencies such as weekly, monthly and quarterly are also valid.

5.1.2 Description of Relation

In the previous sections; the characteristics of the constructs associated with the AIM programme were described. In this section, the relationships that exist between these constructs are described.

Regarding the PMP, there is a hierarchy of relationships involving the key distinctions that make up the functional schema. The first relationship is between PROGRAMME GOAL and PROGRAMME

OBJECTIVE. This relationship is such that each goal is associated with multiple objectives. Conversely, every objective must belong to only a single goal. The second relationship is between programme objective and programme outcome. This relationship is such that there are several outcomes that contribute towards the achievement of a single objective. But a single outcome may contribute towards the achievement of multiple objectives.

In this section various distinctions, attributes and relations that constituted the AIM programme were identified and described. Overall, AIM was made up of the following distinctions: PROGRAMME GOAL; PROGRAMME OBJECTIVE; PROGRAMME ACTIVITY; PROGRAMME PROVIDER; PROGRAMME ADMINISTRATOR; PROGRAMME FUNDER; TARGET POPULATION; PROGRAMME OUTCOME; and PROGRAMME MANAGEMENT PLAN. These distinctions have been described in terms of their essential characteristics and attributes. In doing so, it is possible to identify salient issues that had implication for the design of the conceptual model proposed in the study:

- AIM defined five PROGRAMME OBJECTIVES. Within these five objectives, three categories of beneficiaries were targeted: institutions within the programme districts that were responsible for delivering HIV/AIDS services; people affected/infected with HIV/AIDS; and the entire population of the programme districts. The key issues that were to be addressed included institutional capacity building in relation to HIV/AIDS programming and coordination; access to and utilization of quality HIV/AIDS clinical and prevention services. However, actual beneficiaries that were identified from the various AIM reports included the District Aids Committees (DACs); Youth and Adult; and HIV+ pregnant women. It is likely that there was a gradual shift in targeting as the AIM programme evolved either in response to external influence such as change in funding or as a result of a general shift in programming focus. These issues were essential in conceptualizing the model developed in this thesis.
- AIM was delivered through a system of grant to national and local-level organizations. This arrangement portrayed a hierarchical mode of programme delivery that had to be taken into consideration when developing the conceptual model proposed in the study.
- Programme administration was clustered within two levels: at the higher level, AIM was administered through a consortium comprising of JSI Research Training Institute, Inc., World Education and World Learning with JSI acting as a lead consortium member. At the lower level; AIM was administered through District HIV/AIDS Committee (DAC) and the JSI national-level office. This arrangement displayed a hierarchical and multi-layered setup of programme administration.
- Funding for AIM evolved considerably throughout the programme life-time. Initially funding came from USAID and CDC, but later additional funding started to come from LIFE Initiative

and USAID infectious disease resources for tuberculosis. However at end of programme, AIM was receiving only PEPFAR funding.

5.2 The Northern Uganda Malaria AIDS Tuberculosis Programme (NUMAT)

The Northern Uganda Malaria AIDS & Tuberculosis (NUMAT) was an ongoing five-year USAID-funded programme that began in August 2006 with the goal of expanding access to and utilization of HIV, tuberculosis, and malaria prevention, treatment, care and support activities in the Northern Uganda districts of Amolatar, Amuru, Apac, Dokolo, Gulu, Kitgum, Lira, Oyam and Pader. NUMAT planned to expand the geographic coverage and populations served through strengthening local government responses, expanding the role of communities in planning, implementation and monitoring activities, and building upon existing networks. NUMAT is implemented by the JSI Research & Training Institute, Inc. in partnership with the AIDS Information Centre (AIC), World Vision, local governments, and civil society partners.

5.2.1 Description of distinction

Programme Goal

NUMAT has one programme goal: expanding access to and utilization of HIV, tuberculosis, and malaria prevention, treatment, care and support activities in the Northern Uganda districts of Amolatar, Amuru, Apac, Dokolo, Gulu, Kitgum, Lira, Oyam and Pader. Through an analysis of this goal statement, the following attributes were derived:

- Goal-statement: this attribute represents the text used to describe the programme goal. For example; expanding access to and utilization of HIV, tuberculosis, and malaria prevention, treatment, care and support activities in the Northern Uganda districts of Amolatar, Amuru, Apac, Dokolo, Gulu, Kitgum, Lira, Oyam and Pader;
- What-to-improve: this attribute describes the situation that NUMAT intended to make better through its interventions. For example; **access to** HIV, tuberculosis, and malaria prevention, treatment, care and support activities; **utilization of** HIV, tuberculosis, and malaria prevention, treatment, care and support activities.
- Who-benefits: this attribute describes the person or entity that was expected to gain from the improvement being sought. NUMAT targeted pregnant women, particularly those testing positive for HIV; HIV Positive individuals; the most-at-risk-populations (MARPs); TB patients; malaria patients and the community members within the programme districts of Amolatar, Amuru, Apac, Dokolo, Gulu, Kitgum, Lira, Oyam and Pader;

Programme Objective

NUMAT contains five programme objectives: (1) Improve coordination of HIV, TB, and malaria responses with emphasis on district, sub-county, and internally-displaced persons at camp levels; (2) Increase access to and use of quality HIV, TB, and malaria prevention, care, and treatment services; (3) Decrease vulnerabilities for specific groups to HIV and other infectious diseases; (4) Increase access of PLA and their families to other health services through effective partnerships; (5) Improve use of strategic information. The following are the attributes associated with programme objective:

- Objective-statement: this attribute represents a description of a particular objective. For example; Increase access to and use of quality HIV, TB, and malaria prevention, care, and treatment services;
- What-to-improve: this attribute describes the situation that the programme intended to make better through its various interventions. For example; coordination of HIV, TB, and malaria responses; access to and use of quality HIV, TB, and malaria prevention, care, and treatment services;
- Who-benefits: this attribute describes the person or entity that is expected to profit from the improvement being sought. An objective may benefit one or more persons or entity. For example; people living with HIV (PLA), family of people living with HIV;

Programme Activity

This category or construct describes the tasks that NUMAT executed during its life-time. The extracts that are presented here were mainly derived from the annual and the mid-term evaluation reports. These reports, particularly the annual report, describe activities that have been implemented within a 12-month period. The key attributes that have been derived for this construct are:

- Service-Area: this attribute describes the broad area of service or package that NUMAT dispensed to its beneficiaries. NUMAT defined several service area including: HIV Counseling and Testing (HCT); Prevention of Mother to Child Transmission of HIV (PMTCT); Palliative care; CB-DOTS & TB/HIV Collaboration activities; Anti-Retroviral Therapy (ART); Laboratory services; Human Resources for health; Malaria services; HIV Prevention for the Youth; HIV Prevention for other Adults; Sexual-and Gender-based violence (SGBV); Stigma and Discrimination.
- What-was-done: this attribute describes the actions that NUMAT performed during a given reporting period. The description was provided in narrative format, and would sometimes also embody statistics on the quantity of service or product produced. The description may also

include a description of the approach used to deliver the service, along with the population reached and rationale behind the activity. An excerpt from annual reports is presented below:

NUMAT distributed IEC/BCC materials targeting the youth, people with disabilities and the general population so as to create awareness on availability of HCT Services

- Output: this attribute describes the products emanating from the activities that were carried out. The output reflects progress on each particular service area. For example, under the HCT service area; the following output was reported under the HCT service area:

127,239 clients tested for HIV at 91 facilities

- Challenge: this attribute describes any impediment encountered during activity implementation. Example from the 2007-2008 annual report:

NUMAT still faced challenges affecting HCT service delivery including delays in supply of HIV test kits from the national medical stores (NMS) to the HCT sites

- Remedy: an attribute that describes the workaround proposed or used to counter the challenges encountered during activity implementation. Example from the 2007-2008 annual report:

NUMAT and AIC averted test kit shortage by providing buffer supplies to HCT sites. Additionally, sites were supported to submit timely logistics orders and reports to supply chain management systems (SCMS) projects at NMS

- Time-Period: this attribute specifies the frequency with which NUMAT was required to submit activity implementation report. For NUMAT, activity reports were submitted on an annual basis. But reporting frequency may be set to any time period including weekly; monthly; and quarterly.
- Success-story: an attribute that represents a short, free-standing tale about an interesting or remarkable incident or situation that occurred or was observed. An excerpt from the 2008-2009 annual report is provided below:

Community Members Caring for People with HIV&AIDS: The Story of Charles Odong

When NUMAT initiated its home-based care (HBC) program in Amuru District, many PHAs volunteered to be trained as home visitors because they felt they could better understand the needs of their peers. For Charles Odong, a farmer, husband and father of three children, the motivation was different. Although Charles is not a PHA himself, he has always had a passion to care for PHAs; now, as a trained home visitor, Charles cares for PHAs in Kal Parish, Atiak Sub-County. Charles has recently left the IDP camp to resettle in his village, Parabel, where he continues his HBC efforts.

About his role as a home visitor, Charles said: “To date, I have supported 42 PHAs. It gives me joy to see each of them back on their feet as a result of the comprehensive care and treatment services they receive. I am proud to be part of the care team.”

Source: NUMAT 2008-2009 annual progress report

Programme Provider

The NUMAT programme was delivered by various categories of providers including staff of local government and community-based entities. One of NUMAT’s core strategies was strengthening existing service provision sites to increase coverage and quality of services offered. To that end, NUMAT supported health facilities to deliver quality malaria, AIDS and Tuberculosis services within their catchments. In effect, health facilities (and their staff) are one category of programme providers. At district level, several entities such as the District AIDS Committee (DAC) and several departments at local government level were involved in direct service delivery. This constitutes another category of programme provider. Besides local government entities, civil society organisations at district and community level were also involved in service delivery. Organisations of people living with HIV/AIDS were also supported to help them extend services to their colleagues in remote areas. Below are excerpts from NUMAT annual reports:

NUMAT has also supported institutional and organizational development of Civil Society

Organizations and Faith Based Organizations (FBOs) involved in MAT responses through sub-granting and undertaking of organizational self-assessment and strategic plan development, especially for PHA networks for efficient and effective delivery of services to their constituencies

NUMAT has built the technical and management capacity of districts to deliver organized, high-quality services, to reduce duplication among and between districts and among other implementers and to support coordinating and accountability strategies by funding district-designed activities.

Through a critical assessment, the following attributes were derived:

- Provider-name: an attribute that may represent a unique identifying characteristic of a programme provider.
- Provider-type: an attribute that may represent the kind of organization that each provider belongs. Recognized types include CBO; FBO; NGO; PSO; and Government;

Programme Administrator

NUMAT is implemented by the John Snow Inc. (JSI) Research and Training Institute, Inc. in collaboration with the AIDS Information Centre (AID) and World Vision. A field office for the programme was established in the northern Uganda district of Gulu. Major programme coordination and support activities such as allocation of resources, recruitment, training and supervision of programme providers were executed by the field office. Through a critical assessment, the following attributes were derived:

- Implementing-Agency: an attribute that specifies the name of the main organization or entity responsible for implementing the NUMAT programme. E.g. JSI Research Training Institute.
- Implementing-Partner: an attribute that specifies the names of all other organizations or entities supporting the implementing agency. E.g. World Vision.

Programme Funder

The sole funder of NUMAT was the United States Agency for International Development (USAID) office in Uganda. The programme was funded through cooperative agreement number: 617-A-00-06-00090-00. This information was provided on the second page of every progress report:

The Northern Uganda Malaria, AIDS & Tuberculosis Programme (NUMAT), implemented through Agreement No: 617-A-00-06-00090-00 is funded by the United States Agency for International Development.

The attributes associated with the distinction are:

- Funder-name: an attribute that may represent a unique identifying characteristic of the entity financing the programme. For example USAID;
- Funding-amount: an attribute that may represent the sum of money that the Funder is committing or making available for the programme. For example; \$120m.

- Funding-period: an attribute that may represent the time period within which the fund was made available; e.g. 2002-2007.

Target Population

The NUMAT programme targeted different categories of people from across the nine programme districts. The population targeted by NUMAT included: Pregnant Women, particularly those testing positive for HIV; HIV Positive individuals; the most-at-risk-populations (MARPs); TB patients; malaria patients and the community members within the programme districts. NUMAT offered specific services across the continuum of prevention, treatment and care to each target group.

The attributes associated with the distinction are:

- Target-population: an attribute that represents a unique identifying characteristic of the people or entity targeted by the programme. For example Youth; Adult; HIV pregnant women.
- Service-provided: an attribute that specifies the products that the programme provides to each target population. For example Prevention of Mother to Child Transmission.

Programme Outcome

The annual progress and the mid-term evaluation reports did not discuss issues of outcomes. As such, no attributes were derived for this distinction.

5.2.2 Description of Relation

There are a number of relationships observed among the constructs that have been discussed above. The first relationship is between PROGRAMME OBJECTIVE and PROGRAMME ACTIVITY. This relationship can be viewed from two positions. The first position describes the fact that NUMAT designed and executed several activities under each programme objective. Therefore, a PROGRAMME OBJECTIVE may associate with one or several PROGRAMME ACTIVITY. The second position describes the fact that NUMAT designed and implemented certain activities to address or contribute towards realization of one or more objectives. Therefore, a PROGRAMME ACTIVITY may associate with one or more PROGRAMME OBJECTIVE.

A second relationship exists between PROGRAMME ACTIVITY and PROGRAMME PROVIDER. This relationship is also viewed from two positions. The first position describes the fact that organizations that participated in implementing NUMAT where each responsible for implanting several activities. In effect, a single PROGRAMME PROVIDER can implement one or more

PROGRAMME ACTIVITY. The second position describes the fact NUMAT standardized its service delivery and in as much as there were several grantees responsible for implementation; these grantees would sometimes carry out similar activities. Therefore, A PROGRAMME ACTIVITY can be implemented by one or more PROGRAMME PROVIDER.

In this section various distinctions, attributes and relations that constituted the NUMAT programme were identified and described. Overall, NUMAT was made up of the following distinctions: PROGRAMME GOAL; PROGRAMME OBJECTIVE; PROGRAMME ACTIVITY; PROGRAMME PROVIDER; PROGRAMME ADMINISTRATOR; PROGRAMME FUNDER; and TARGET POPULATION. These distinctions have been described in terms of their essential characteristics and attributes. In doing so, it was possible to identify salient issues that had implication for the design of the conceptual model proposed in the study:

- NUMAT was delivered through a partnership arrangement with district-level organizations. These organizations were provided with funds to implement agreed activities within a specified period of time.
- Programme administration was a collaborative effort involving JSI Research Training Institute, Inc., AIDS Information Centre (AID) and World Vision with NUMAT being the led agency. At the lower level; NUMAT established a field-level office where programme administrative issues are addressed.

5.3 The President's Emergency Plan for AIDS Relief (PEPFAR)

The United States President's Emergency Plan for AIDS Relief (PEPFAR) was launched in 2003 to combat global HIV/AIDS. PEPFAR's first phase ended in 2008, and following a re-authorization in 2008 (with a doubling of funding), PEPFAR entered into its second phase (that was running until 2013). Uganda is one of PEPFAR's 15 focus countries that received funds to support a comprehensive HIV/AIDS prevention, care, and treatment programme. The programme is implemented through a system of grants awarded to country-wide network of implementing partners (IPs), most of which are international nongovernmental organizations (INGOs). In-country programme management is shared among the in-country United States Government (USG) agencies, notably United States Agency for International Development (USAID), Center for Disease Control (CDC), National Institute of Health (NIH), Department of Defense (DOD), Peace Corp, Department of Labor and United States (U.S) Embassy.

5.3.1 Description of distinction

Programme Goal

PEPFAR planned to work in partnership with host nations to support the realization of six legislative goals in the areas of treatment, prevention, care and workforce. The six programme goals are: (1) treatment for at least 3 million people; (2) prevention of 12 million new infections; (3) care for 12 million people, including 5 million orphans and vulnerable children (OVCs); (4) 80% coverage of testing and counselling among pregnant women; (5) 80% coverage of ARV prophylaxis for HIV-positive pregnant women; and (6) professional training for 140,000 new health care workers. The Next Generation Indicator reference guideline summarised these legislative goals as shown in the table below.

Table 7: PEPFAR legislative goals and indicator (source: Next generation indicator reference guide)

PEPFAR legislative goal		Monitoring indicator
TREATMENT		
Treatment for at least 3 million people		Percentage of adults and children with advanced HIV infection receiving antiretroviral therapy
Prevention		
12 million new infections averted		No routine monitoring indicator – Goal is measured through modeling at HQ
80% coverage of testing and counseling among pregnant women		Percent of pregnant women with known HIV status (includes women who were tested for HIV and received their results)
80% coverage of ARV prophylaxis for HIV – positive pregnant women		Percent of HIV – positive women who received antiretroviral to reduce the risk of mother – to – child transmission
Care		
Care for 12 million people, including 5 million orphans and vulnerable children		Number of eligible and children provided with a minimum of one care service (disaggregated by age)
Human Resources for Health – Workforce		
Profession training for 140,000 new health care workers		Number of new health care workers who graduated from a pre-service training institution

Source: Next generation indicator reference guide

The legislative goals and targets listed in the table above are the key parameters within which PEPFAR planned and delivered intervention strategies. The following attributes are identified for the PROGRAMME GOAL:

- Goal-statement: this attribute represents the text used to describe the programme goal. For example; prevention of at least 12 million new infections;
- What-to-improve: this attribute describes the situation that the PEPFAR programme intended to make better by intervening. For example; prevention; treatment; care; testing & counselling; professional training.
- Who-benefits: this attribute describes the person or entity that is expected to profit or gain from the improvement being sought. For example; OVC's; health care workers.
- Target: this attribute describes the magnitude or amount of improvement that the programme expected to attain within a specified period of time. Within this attribute a pass mark was also established for the programme. Example; **prevent 12 million** new infections;

Programme Objective

The PEPFAR programme did not define any programme objective. Programming of activities and strategies were guided by the six programme goals discussed above.

Programme Activity

The PEPFAR programme is an ambitious programme implemented across several countries. However, although PEPFAR required each implementing country to submit annual progress report, no such reports could be found on the Internet and only summary progress report with no narrative section could be located. An excerpt of the 2009 annual consolidated report for the prevention programme area is shown in the table below:

Table 8: Country summary for PMTCT legislative goals and indicator

Prevention: FY2009 Prevention of Mother-to-Child Transmission Program Direct Results								
Country	Pregnant women receiving HIV counseling and testing services	Number of pregnant women receiving prophylaxis	HIV + ARV	Estimated infections averted ³	infant	HIV		
Angola	27,400	NA		NA				
Botswana	0	0		0				
Angola	27,400	NA		NA				
Cambodia	48,000	300		57				
Caribbean Regional	NA	NA		NA				
China	121,200	300		57				
Cote d'Ivoire	224,900	7,800		1,482				
Democratic Republic of Congo	98,300	900		171				
Dominican Republic	20,800	100		19				

Ethiopia	436,700	8,300	1,577
Ghana	NA	NA	NA
Guyana	10,500	200	38
Haiti	155,800	1,900	361

Source: 2009 annual report to congress

PEPFAR's documentation of programme activities focused primarily at describing outputs that each country produced against established targets in numerical terms as described in the table above. No description was provided on the activities that produced the reported outputs. Reported outputs were categorized by programme goal and technical area. A technical area refers to a broad strategy along the areas of HIV/AIDS prevention, care, treatment and workforce; which is believed to hold great promise in ameliorating the HIV/AIDS problem situation in that area. Overall, PEPFAR defined 14 technical areas. Under each technical area, reporting of outputs was done using standardized indicators. Each technical area defined a set of indicators that are applicable to the technical area. For example, the outputs reported in table 4 above correspond to the prevention of mother-to-child transmission (PMTCT) technical area. The two outputs are: number of pregnant women receiving HIV counselling and testing and number of HIV+ pregnant women receiving ARV prophylaxis.

Through an analysis of the goal statement, the following attributes were derived:

- Goal-statement – a mutual property that specifies the programme goal that the activity contributes towards. The programme goal may be any one of the six goals defined for the PEPFAR programme.
- Technical-area – an attribute that specifies one of the fourteen broad strategies in the areas of HIV/AIDS prevention, care, treatment and workforce that this activity contributes towards. For example, prevention of mother-to-child transmission (PMTCT).
- Output – this attribute describes the products and services resulting from the activities that have been carried out. Outputs are described in quantitative terms and using the pre-

specified indicators such as number of pregnant women receiving HIV counselling and testing services; number of HIV+ pregnant women receiving ARV prophylaxis.

- Time-Period: this attribute specifies the frequency with which PEPFAR was required to submit activity implementation report. For PEPFAR, activity reports were submitted on an annual basis. But reporting frequency may be set to any time period including weekly; monthly; and quarterly.

Programme provider

PEPFAR was implemented through partnership arrangements with governments, non-governmental organizations including faith- and community-based organisations and the private sector of the host country. A PEPFAR partner can be one of two types: a prime partner or a sub partner. The COP guideline defines a prime partner as an organization that receives funding directly from, and has a direct legal relationship (contract, cooperative agreement, grant, etc.) with an USG agency. It defines a sub partner as an entity that receives a sub-award from a prime partner or another sub-partner under an award of financial assistance or contract and is accountable to the prime partner or other sub-partner for the use of the Federal funds provided by the sub-award or sub-contract.

Two attributes associated with this distinction are identified:

- Partner-name – an attribute that represents a unique identifying characteristic of a programme provider. This attribute captures the names of both prime and sub partners.
- Partner-type – an attribute that may represent a category of organization that a provider may belong. Each programme provider can only belong to one type of organisation. A list of possible organisation types from the COP guideline includes: FBO; NGO; host country government agency; private contractor; university; multi-lateral agency; other USG agency; own agency and parastatal.

Programme Administrator

Owing to its nature, the administration of PEPFAR was found to be hierarchical. At the global level, the mandate to manage and coordinate PEPFAR programmes in developing countries is vested in the Office of the Global AIDS Coordinator (OGAC). OGAC is accountable to the U.S congress for all successes and failures of PEPFAR. At country level, PEPFAR is administered by the US ambassador who reports to OGAC and has the responsibility to sign-off on all planning and reporting documents that are submitted to OGAC. At country level, the in-country USG agencies, notably USAID, CDC, NIH, DOD, Peace Corp, Department of Labor and U.S Embassy are responsible for supporting programmes among their respective implementing partners/constituency and collaborating to plan

programming at the beginning of each fiscal year and for periodic results. Therefore, in each focus country; PEPFAR was administered by the in-country USG agencies. The following are the attributes that have been derived for this distinction:

- Implementing-Agency: an attribute that specifies the name of the USG agency responsible for specified PEPFAR interventions (also known in PEPFAR lingua as Funding Mechanisms). E.g. USAID; CDC.
- Implementing-Partner: an attribute that specifies the names of the prime partner who is responsible for conducting the activities of specified PEPFAR interventions (funding mechanisms). E.g. PLAN International; World Vision.

Programme Funder

PEPFAR is funded by the United States government through the Office of the Global AIDS Coordinator (OGAC). But at country level, funding was provided through the in-country USG agencies - notably USAID, CDC, NIH, DOD, Peace Corp, Department of Labor and U.S Embassy. It was noted that Prime Partners were funded by the in-country USG agencies. The following attributes have been derived:

- Funder-name: an attribute that represents a unique identifying characteristic of the entity that is financing a particular PEPFAR intervention (also known as funding mechanism). For example USAID;
- Funding-amount: an attribute that may represent the sum of money that the Funder is committing or making available for funding interventions (funding mechanisms) within its jurisdiction; e.g. \$120m;
- Funding-period: an attribute that represents the time period within which the fund was made available; e.g. 2009.

Target Population

PEPFAR holds a unique place in the history of public health for its size and scope. Unlike many programmes that tended to treat prevention, treatment and care individually and in isolation, PEPFAR has moved towards an integrated prevention, treatment and care for HIV/AIDS. As such, PEPFAR targets several categories of beneficiaries. Reading through the list of PEPFAR indicators, there appears to be five different kinds of beneficiaries. A first major category of beneficiaries is the entire inhabitants of Uganda. This has been targeted with HIV/AIDS prevention, testing and counselling programming. A second category is HIV+ pregnant women. This has been the target of prevention of mother to child transmission programming. A third category is the Ugandan male population,

particularly the un-circumcised male. This category has been the target of the male circumcision programming. A fourth category is HIV+ individuals. This category has been the target of the prevention, treatment and care programming. A fifth category is health workers, which were targeted mainly for health system strengthening programming.

The following attributes have been identified:

- Target-population: an attribute that represents a unique identifying characteristic of the people or entity targeted by the programme. For example HIV+ individuals;
- Service-portfolio: an attribute that specifies the various services or products that the programme provides to each target population. For example ART; HCT.

Programme Outcome

Assessment of the various PEPFAR documents did not result in the identification of any outcomes. In fact, PEPFAR relied mainly on a comprehensive list of indicators defined in the next generation indicator reference guide. This indicator list provided a framework within which PEPFAR assessed and reported progress towards programme goals. As such, it was not feasible to discuss attributes associated with this distinction.

Country Operational Plan

The Country Operational Plan (COP) constitutes the key annual planning resource for PEPFAR. It is the vehicle for documenting USG annual investments and anticipated results in HIV/AIDS, and the basis for approval of annual USG bilateral HIV/AIDS funding. It serves as the annual work plan for the in-country PEPFAR programme. The following attributes have been identified:

- Operating Unit¹⁷ background - an attribute that describes the overall operating unit context, key indicators, state of the overall health sector, financing situation and status of the national HIV/AIDS strategy. This is a qualitative or narrative description.
- Partnership Framework¹⁸ - this attribute captures key information from the partnership framework, if it has been signed with the host country. Key information captured include the goal, objective, output and outcome of the Partnership Framework; financial commitment of key partners; the areas of focus of the financial commitment for each partner; and commitment to specified activities for each partner.

¹⁷ In PEPFAR, this terminology loosely refers to a particular country that is implementing PEPFAR. It is a preferred terminology because it also applies to regional and headquarters organizations.

¹⁸ Presents an integrated approach to combating HIV/AIDS epidemic that involves developing a joint five-year strategic framework for cooperation between USG, the partner government, and, in some cases, other partners.

- National Level Indicator Target – this attribute describes the expected achievements of all contributors to a country's HIV programme including the host country government and all of its stakeholders, donors, and civil society organizations. The national level indicators are defined to show achievements along the continuum of output; outcome and impact.
- Technical Area– this attribute describes a broad strategy along the areas of HIV/AIDS prevention, care, treatment and workforce that is believed to hold great promise in ameliorating the HIV/AIDS problem situation in the particular area. PEPFAR defined 14 technical areas including PMTCT; Sexual Prevention; Biomedical Prevention; Counseling and Testing; Adult Care and Treatment; Pediatric Care and Treatment; TB/HIV; OVC; ARV Drugs; Laboratory Infrastructure; Strategic Information; Health Systems Strengthening; Human Resources for Health; Gender.
- Technical Area Narrative – this attribute provides an overview of the country's strategy in the various areas of prevention, care, and treatment. For example, under the technical area of PMTCT; the following is an extract from the 2007 COP:

In FY07, the USG will continue to work with the MOH and other partners to support the implementation of Uganda's new, revised national PMTCT policy. Routine opt-out CT for pregnant women will be strengthened and expanded in all USG supported sites. Training materials will be revised to be consistent with the new national PMTCT policy. Refresher training for trainers and service providers at all sites will be conducted. In addition, midwives will be trained to run rapid HIV tests in ANC. Follow-up training for Nursing and Clinical Officers schools' tutors will be conducted to support the integration of PMTCT into pre-service education of medical and paramedical staff. PMTCT services will be expanded to select HC IIIs to increase access of services in the rural areas where the majority of the population live. Other strategies to improve uptake and coverage of PMTCT services will be expanded through peripheral health units conducting outreaches and introducing routine intrapartum CT

- Technical Area Budget – this attribute denotes the total amount of money required to implement the strategies specified under the technical area narrative.
- Provider Budget - this attribute denotes the amount of money that is budgeted for each entity or organization implementing activities under the technical area narrative. Each technical area is implemented by several providers.
- Provider Activity Narrative - this attribute describes the activities that each provider implementing activities under a specified technical area intends to undertake. Each technical area is implemented by several providers.

- Provider Indicator Target – this attribute describes the expected achievements of a single provider along the continuum of output; outcome and impact. Each technical area is implemented by several providers; with each provider having its own indicator target.
- Timeframe – this attribute describes the time period covered by the COP. It indicates the amount of time that is required to complete planned activities and to realize National-level or programmatic targets. COPS timeframe covered a 12-month period e.g., January 2009 – December 2009 or October 2009 – September 2010.

Indicator

Indicator describes major dimensions and standards used for measuring/tracking progress of PEPFAR. In the Nextgen indicator reference guide, several attributes are defined for an indicator. Nine of the attributes relevant to the thesis are described below.

- Indicator-Name: an attribute that may reflect actual dimension of performance being measured. E.g. pregnant women with known HIV status (includes women who were tested for HIV and received their results); eligible adults and children provided with a minimum of one care service.
- Type of Indicator (Reporting level): an attribute that classifies “indicator” according to levels. The two levels are direct and national. Direct-level indicators measure results that are accredited to PEPFAR only. National-level indicators measure results that PEPFAR only makes partial contribution towards, with several other programmes also contributing to its realization.
- Type of Indicator (degree of Importance): an attribute that specifies the significance of an indicator in tracking progress in HIV/AIDS programming. PEPFAR recognizes three degrees of importance: essential/reported to HQ; essential/not reported to HQ and recommended.
- Type of Indicator (M&E result hierarchy): an attribute that classifies indicator according to the standard M&E results classification. PEPFAR recognizes three types of results: output, outcome and impact.
- Numerator: an attribute that specifies the expression used as a numerator in the mathematical procedure for calculating the value of an indicator. The expression may or may not refer to an indicator.
- Denominator: an attribute that specifies the expression used as a denominator in the mathematical procedure for calculating the value of an indicator. The expression may or may not refer to an indicator.
- Disaggregation: an attribute that specifies how results of an indicator may be broken down. Disaggregation types include sex; age; and most-at-risk population group.

- Data collection frequency: an attribute that specifies how often data on the indicator shall be collected, summarized and reported. Various options for the attribute exist including daily, weekly, monthly, quarterly and annually. Each indicator can specify only one data collection frequency.

5.3.2 Description of Relation

This section discusses the distinctions, functional schemas and attributes associated with the PEPFAR programme. It was noted that PEPFAR used statistical summaries to report progress and results. The statistics are reported within a framework of indicators, which defined the various performance dimensions against which programme performance and results were measured. The following are the key relationship observed.

- A single indicator such as *number of pregnant women receiving HIV counseling and testing* may be associated with several target (target result) and actual (actual result) values. This comes about because each programme provider is required to establish target for each indicator on an annual basis. Likewise, each programme provider must report on the actual achievement in relation to each indicator annually.
- During a reporting cycle, a programme provider may associate with only a single programme administrator (e.g. USAID, CDC). However, a programme administrator may associate with several providers during a particular planning/reporting cycle.

This section identifies and describes the various distinctions, attributes and relations that constitute the NUMAT programme. Overall, PEPFAR was made up of the following distinctions: PROGRAMME GOAL; PROGRAMME ACTIVITY; PROGRAMME PROVIDER; PROGRAMME ADMINISTRATOR; PROGRAMME FUNDER; TARGET POPULATION; COUNTRY OPERATING PLAN; and INDICATOR. These distinctions have been described in terms of their essential characteristics and attributes. In doing so, it was possible to identify salient issues that had implication for the design of the conceptual model proposed in the study:

- PEPFAR was delivered through a system of grant awarded to several sub-partners implementing partners (who themselves received grant from an in-country USG agency) or another sub-partner. The sub-partner is under obligation to use the grant to implement agreed activities.
- Programme administration was multi-layered: at the global level, PEPFAR is administered by the Office of the Global AIDS Coordinator (OGAC). At country level, coordination of PEPFAR programme is performed by the US ambassador. The in-country USG agencies, notably

USAID, CDC, NIH, DOD, Peace Corp, Department of Labor and U.S Embassy are responsible for supporting programmes among their respective implementing partners and collaborating to plan programming at the beginning of each fiscal year and for periodic results. Therefore, programme administration involved four separate types of entities: OGAC; US Ambassador; in-country USG Agencies and Implementing Partners (IP).

5.4 Concluding comments

This chapter identifies and describes various distinctions, attributes and relations for the AIM, NUMAT and PEPFAR case study programmes as well as salient issues in the distinctions under each case study programme and how they impact on the development of the conceptual model proposed in the thesis. In this concluding section, observed similarities and differences in the distinctions, attributes and relations across the three case study programmes are discussed with a summary for the complete list of distinctions found across the three case study programmes in the table below. A tick under each programme name indicates that the corresponding distinction exists in the programme.

Table 9: Summary of distinctions across AIM, NUMAT and PEPFAR

DISTINCTION	AIM	NUMAT	PEPFAR
Programme goal	√	√	√
Programme objective	√	√	
Programme activity	√	√	√
Programme provider	√	√	√
Programme administrator	√	√	√
Programme funder	√	√	√
Target population	√	√	√
Programme outcome	√		
Programme management plan	√		
Country operational plan			√
Indicator			√

It can be observed that:

- There are distinctions that exist across the three case study programmes. These distinctions include PROGRAMME GOAL; PROGRAMME ACTIVITY; PROGRAMME PROVIDER; PROGRAMME ADMINISTRATOR; PROGRAMME FUNDER; and TARGET POPULATION. The characteristics or attributes of these distinctions were also found to be similar across the three case study programmes, except for PROGRAMME GOAL and PROGRAMME ACTIVITY. This variation was mainly with PEPFAR. For PROGRAMME GOAL, PEPFAR defined an additional attribute for recording the target established for each programme goal. For PROGRAMME ACTIVITY; there were two areas of variation: (1) PEPFAR defined a unique attribute that links an activity to programme goal. (2) Attributes to capture narrative information about an activity such as description of what was done, adjustments made, reasons for the adjustments, challenges encountered, remedy to existing challenges and success story were not defined. These similarities and differences were considered in developing the conceptual model proposed in this thesis.
- The three case study programmes captured very little information on programme inputs or resources. In fact, only PROGRAMME FUNDER and PROGRAMME PROVIDER captured some little information on inputs and use of resources.
- Information relating to programme implementation was the most captured across the three case study programmes. Several of the distinctions discussed such as PROGRAMME ACTIVITY; PROGRAMME MANAGEMENT PLAN; ACTIVITY IMPLEMENTATION and COUNTRY OPERATIONAL PLAN defined a number of attributes that capture information about implementation.

Overall, the distinctions, attributes and relations identified across the three case study programmes provided sufficient input for the development of the conceptual model proposed in this study. Where there were differences across distinctions, such differences were analyzed and transformed into generic requirements.

Chapter 6: Discussion

This chapter presents a first version of a conceptual model for a programme monitoring and evaluation information system. The proposed model has been “baptised” *Reflector* to better highlight the view that it “mirrors” both the practice and theory of M&E. It has been arrived at following a synthesis of insights acquired from review of the literature (chapter 2); the review of systems thinking and conceptual modelling theories (chapter 3); and the analysis of the practice of monitoring and evaluation as documented by three case study programmes (chapter 5).

As pointed out in chapter 4; the question of the thesis is: *How can the key concepts used in programme monitoring and evaluation be arranged to enable the development of adaptable programme monitoring and evaluation software?* The answer to this question is provided in the form of the conceptual model proposed in this chapter. The proposed conceptual model integrates and organizes the key concepts discussed in chapter 5 into coherent and logical patterns that reflect both practice and theory. The process and thinking that lead to the development of the *Reflector* model is the subject of discussion in this chapter.

6.1 Programme goal

The discussion in chapter 5 defines the structure of PROGRAMME GOAL. Each of the three case study programmes specifies similar attributes for this distinction. It is therefore possible to summarise the attributes of this distinction as goal-statement, what-to-improve, who-benefits and target.

In view of the discussion in chapter 2 and 5; the goal-attainment/achievement model is the most prevalent type of assessment identified among the three case study programmes. The goal-attainment/achievement assessment is found to be more prominent under the PEPFAR programme than AIM and NUMAT. In fact, the main issues discussed in the PEPFAR annual reports are related to progress towards attainment of established goals and/or targets. Progress is reported using indicators that had been defined for each programme goal. These indicators specify what parameters shall measure performance on each defined goal as depicted in the table below. Assessment of this distinction is derived mainly from the PEPFAR case study.

Table 10: PEPFAR legislative goals and indicator

PEPFAR legislative goal	Monitoring indicator
Treatment	
Treatment for at least 3 million people	Percentage of adults and children with advanced HIV infection receiving antiretroviral therapy
Prevention	
12 million new infections averted	No routine monitoring indicator – Goal is measured through modeling at HQ
80% coverage of testing and counseling among pregnant women	Percent of pregnant women with known HIV status (includes women who were tested for HIV and received their results)
80% coverage of ARV prophylaxis for HIV – positive pregnant women	Percent of HIV – positive women who received antiretroviral to reduce the risk of mother – to – child transmission
Care	
Care for 12 million people, including 5 million orphans and vulnerable children	Number of eligible and children provided with a minimum of one care service (disaggregated by age)
Human Resources for Health – Workforce	
Profession training for 140,000 new health care workers	Number of new health care workers who graduated from a pre-service training institution

Source: Next generation indicator reference guide

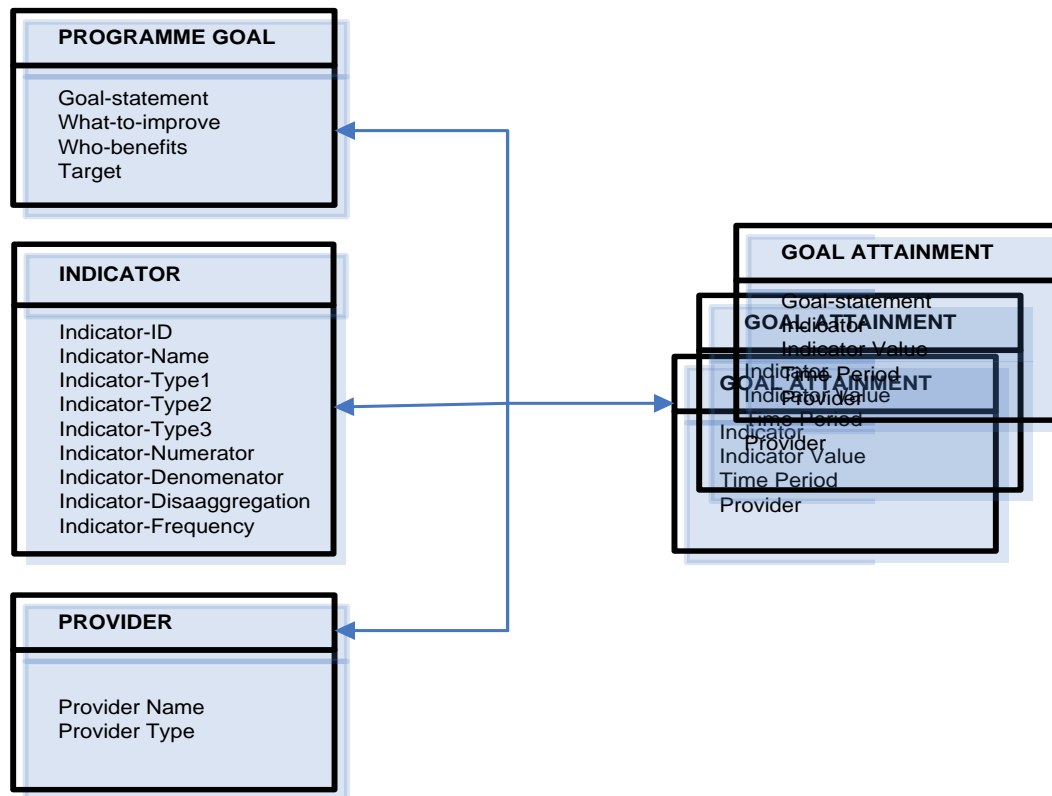
As noted in chapter 5; PEPFAR was implemented through a system of grants to INGOs. During the annual reporting cycle, each grantee (INGO) would provide data on the level of goal and target achievement. This progress was reported using the monitoring indicators specified above. A new distinction named GOAL ATTAINMENT has been defined to capture this information. The following are its attributes:

- ***Indicator:*** an attribute that describes the actual dimension of performance being measured. E.g. *percent of pregnant women with known HIV status (includes women who were tested for HIV and received their results); Percent of adults and children with advanced HIV infection receiving antiretroviral therapy.*

- Indicator Value: an attribute that describes the actual attainment achieved on a specified indicator during a particular reporting period. E.g. *60% of pregnant women with known HIV status (includes women who were tested for HIV and received their results)*;
- Time period: an attribute that represents the range of time within which the result specified above was produced. The most time range has been annual; e.g. 2009.
- Provider: an attribute that represents the name of the entity that produced the results specified above. E.g. Plan International;

In view of the above discussions, it is proposed that the monitoring of programme goal involves the interaction of four key distinctions: PROGRAMME GOAL; INDICATOR; PROGRAMME PROVIDER and GOAL ATTAINMENT. The interaction and relationship among the four distinctions is graphically shown below. As portrayed in the diagram; a single PROVIDER may be associated with many of the information captured in the GOAL ATTAINMENT. Likewise, every unique record in the GOAL ATTAINMENT distinction must related to only a distinct PROVIDER.

Figure 6.1: Conceptual Model for Programme goal



6.2 Programme Objective

The discussion in chapter 5 defines the structure of PROGRAMME OBJECTIVE. From the discussion, it is possible to identify the following three attributes of the distinction: objective-statement, what-to-improve and who-benefits.

In view of the discussion in chapter 2 and 5; the objective-based type of evaluation is the most common assessment identified among the case study programmes. An example of this type of assessment was found in the AIM's end of programme evaluation. For this reason, the discussion in this section draws inspiration from the AIM's end of programme evaluation.

AIM assesses the extent to which it achieves its set objectives through the collection and interpretation of data on the key performance indicators or measures. Under AIM, these performance measures are defined in the so-called programme management plan (PMP). The PMP defines one or more performance measures for every objective. During the end of programme evaluation; the data collected on each performance measure is compared against the targets and baselines established within the PMP. Two new distinctions have been defined to capture this information: OBJECTIVE ATTAINMENT and OBJECTIVE TARGET. Each of the distinction is discussed further below.

The OBJECTIVE ATTAINMENT distinction records and manages the data which the evaluation collects on each performance measure and for each objective. It takes on the following attributes:

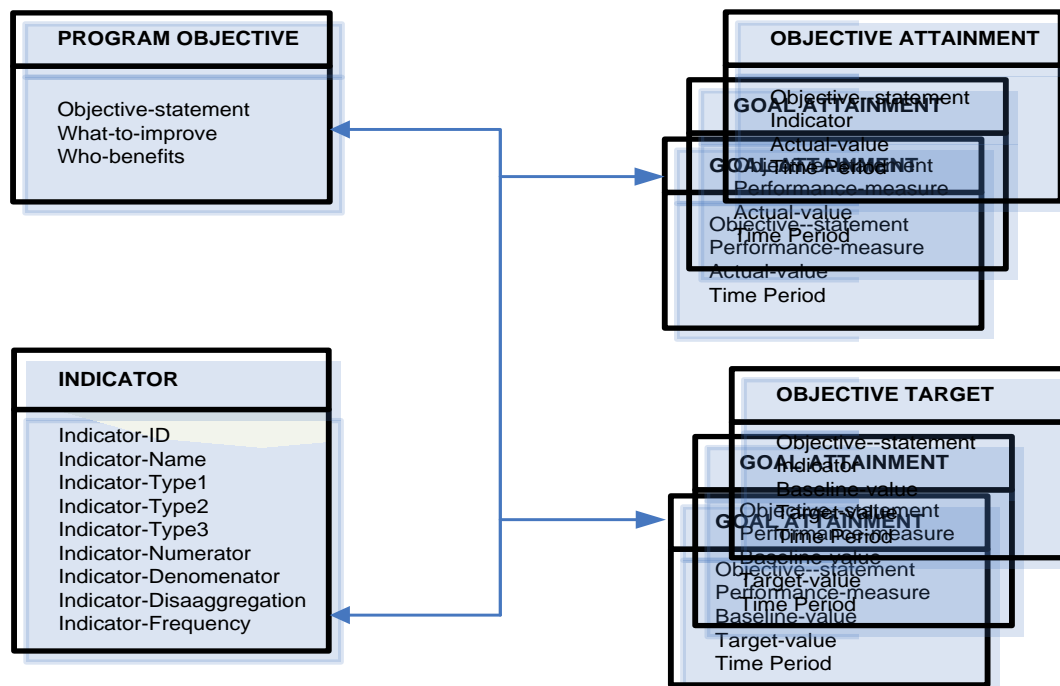
- Objective-statement: this attribute represents a description of a particular objective. For example; *to increase access to and utilization of quality HIV prevention services in selected districts and sub districts.*
- Performance-measure: this attribute defines the indicator that is used to sense existence or non-existence of the change specified in the *what-to-improve* attribute of *programme objective*. For example; *Number of AIM-supported districts that track HIV/AIDS indicators.*
- Actual-value: this attribute specifies the observed increase or decrease in the performance measure and the magnitude by which it has increased or decreased.
- Time-period: this attribute specifies the time period within which the observed increase or decrease in performance measure was produced.

The OBJECTIVE TARGET distinction enables programmes to capture and manage the details of their programme plans in relation to what each objective is expected to achieve and the details of what will be used to measure the achievement. It takes on the following attributes:

- Objective-statement: this attribute represents a description of a particular objective. For example; *to increase access to and utilization of quality HIV prevention services in selected districts and sub districts.*
- Performance-measure: this attribute defines the indicator(s) that is/are used to sense existence or non-existence of the change specified in the *what-to-change*. For example; *Number of AIM-supported districts that track HIV/AIDS indicators.*
- Baseline-value: this attribute specifies the value of the performance measure prior to AIM.
- Target-value: this attribute specifies the expected increase or decrease in the performance measure and by what magnitude.
- Time-Period: this attribute specifies the amount of time that the programme expects to implement activities and to attain the desired results.

In view of the above discussion, it is concluded that monitoring the achievement of objectives involves the interaction of four key distinctions: PROGRAMME OBJECTIVE; INDICATOR; OBJECTIVE ATTAINMENT and OBJECTIVE TARGET. The interaction and relationship among the four distinctions is graphically shown below. As portrayed in the diagram.

Figure 6.2: Conceptual model for programme objective



6.3 Programme Activity

The discussion in chapter 5 defines the structure of programme activity for the three case study programmes. It is observed that each case study programme defines varying attributes for programme activity. A complete list of attributes drawn from across the three case study programmes includes: activity-category, service-area, technical-area, what-was-done, output-produced, adjustment, reason-for-adjustment, challenge, remedy-to-challenge, lesson-learnt, success-story and time-period. Many of the attributes describe issues of programme implementation.

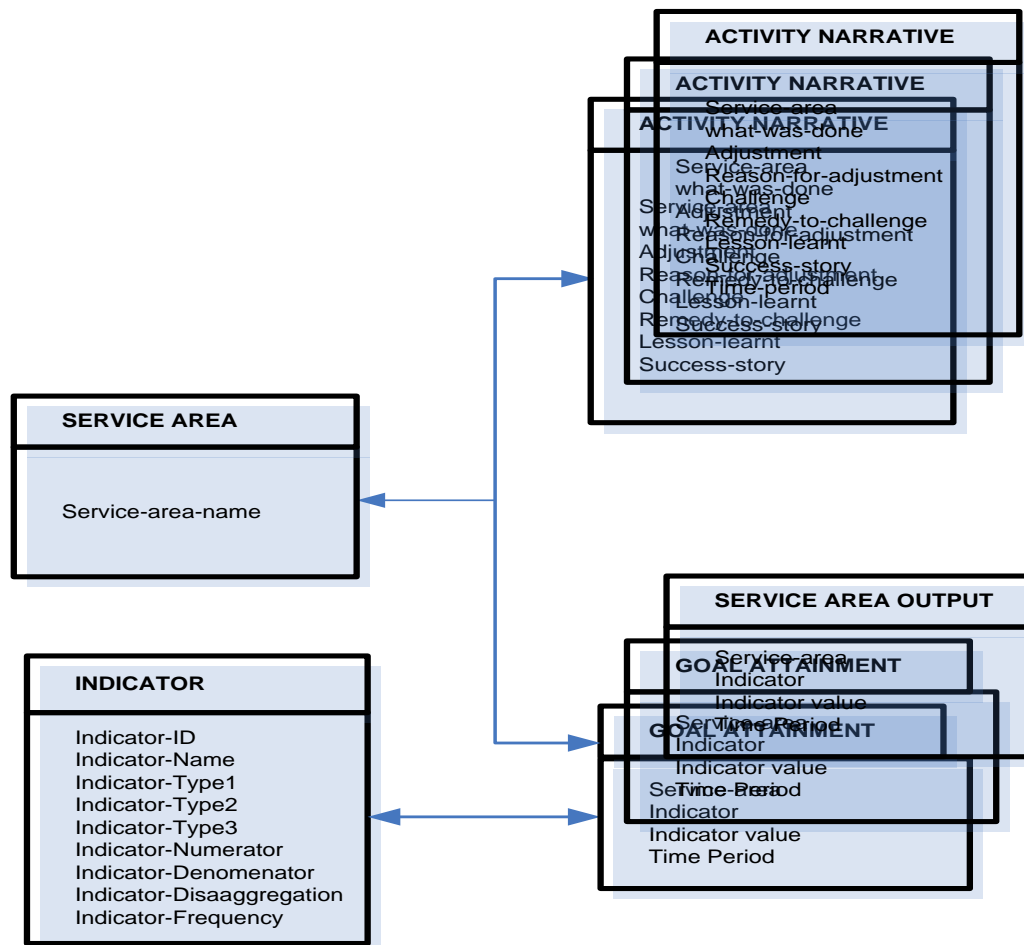
A synthesis of the data from the three case study programmes shows that:

- Each case study programme discusses activities within a form of cluster or grouping. In AIM; activity-category is the attribute used to cluster activities. In NUMAT; service-area is the attribute used to cluster activities. In PEPFAR; technical-area is the attribute used to cluster activities. Each case study programme develops standardized data variables for this attribute. For instance, PEPFAR defines 14 different technical areas under which activities are planned and reported. It was found necessary to store the information on the grouping or clustering of activities into a separate distinction. This distinction has been named SERVICE AREA with just one attribute: service-area-name.
- There are attributes that capture purely qualitative type of information. Such attributes include what-was-done, adjustment, reason-for-adjustment, challenge, remedy-to-challenge, lesson-learnt, success-story. It was found necessary to put all the narrative-like attributes into a separate distinction, which has been named ACTIVITY NARRATIVE.
- There is also one attribute that captures only quantitative type of information. In each case study programme; the quantity of products or services generated by the programme was documented using the output-produced attribute. It was also found necessary to define a separate distinction for managing this kind of information. The distinction has been named SERVICE AREA OUTPUT and contains the following four attributes: service-area; indicator; indicator value and time period. This distinction is dependent on the following two distinctions: SERVICE AREA and INDICATOR. INDICATOR provides a standardized list of data variables within which outputs are reported.

In view of the above discussion, it is possible to conclude that monitoring programme implementation involves the interaction of four key distinctions: SERVICE AREA; ACTIVITY NARRATIVE; SERVICE AREA OUTPUT; and INDICATOR. The relationship among these distinctions is presented in the diagram below. As the diagram portrays; there can be several records for a single SERVICE AREA or

INDICATOR that the ACTIVITY NARRATIVE or SERVICE AREA OUTPUT distinctions may contain at any given time.

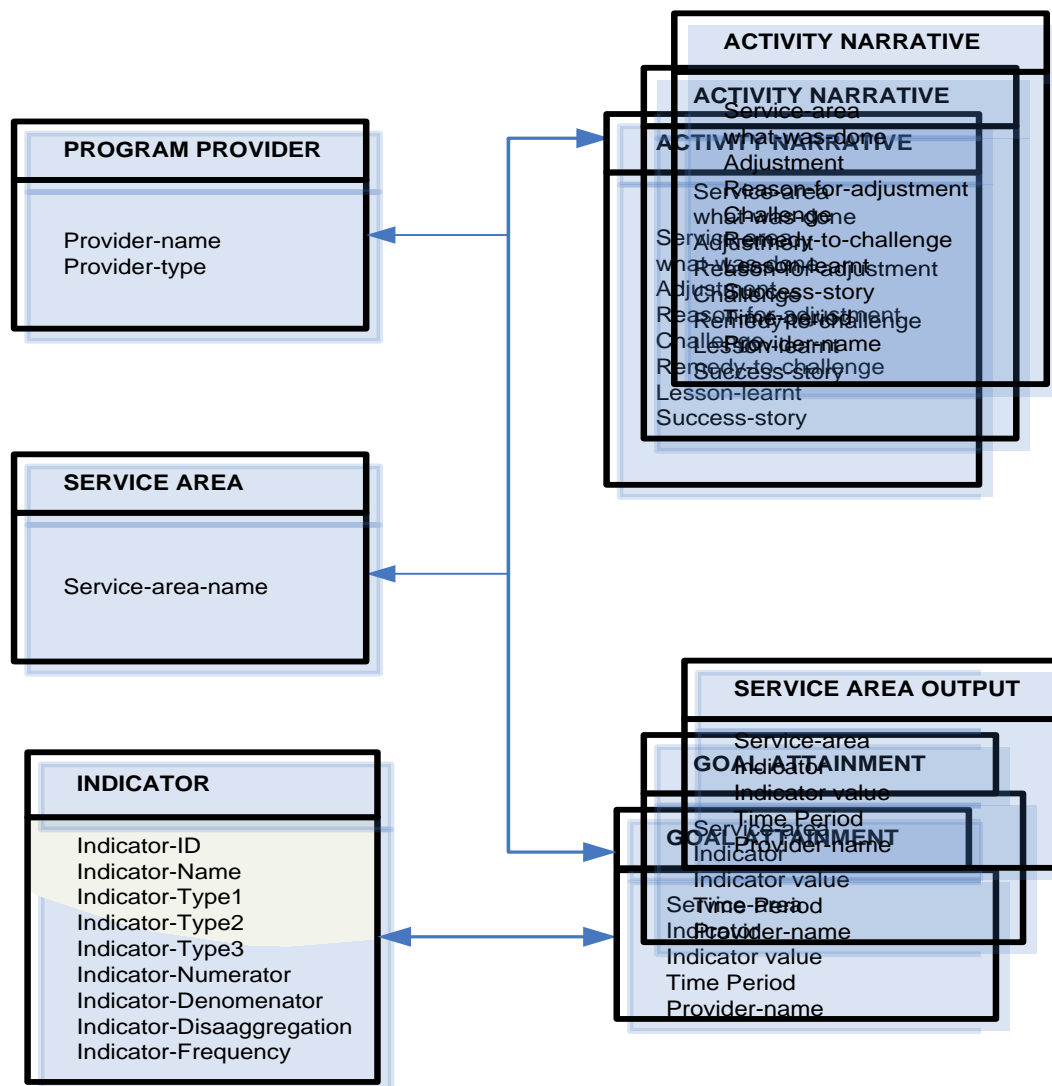
Figure 6.3: Conceptual Model for Programme Activity



6.4 Programme Provider

In the discussion in chapter 5; it is observed that programme implementation can be the mandate of one or more implementing partners. The implementing partners provide periodic accountability for activities that have been conducted and results achieved. In view of this; the conceptualization of PROGRAMME ACTIVITY discussed above was revisited. A major modification has ensured that each record in the ACTIVITY NARRATIVE and SERVICE AREA OUTPUT is related to a distinct PROGRAMME PROVIDER. The modified model is presented in the diagram below.

Figure 6.4: Conceptualizing Programme Provider



Under AIM and PEPFAR, information relating to programme plans is described in PMP and COP respectively. The structures of the distinctions are described in chapter 5. These two distinctions describe the activities that have been planned and the corresponding outputs. For conceptualization, a new distinction named WORKPLAN has been defined to store all planning-related information. The following are its attributes: service-area; what-shall-be-done; what-shall-it-cost; indicator-name; provider-name; time-period.

Figure 6.5: Conceptual Model for Programme Work plan

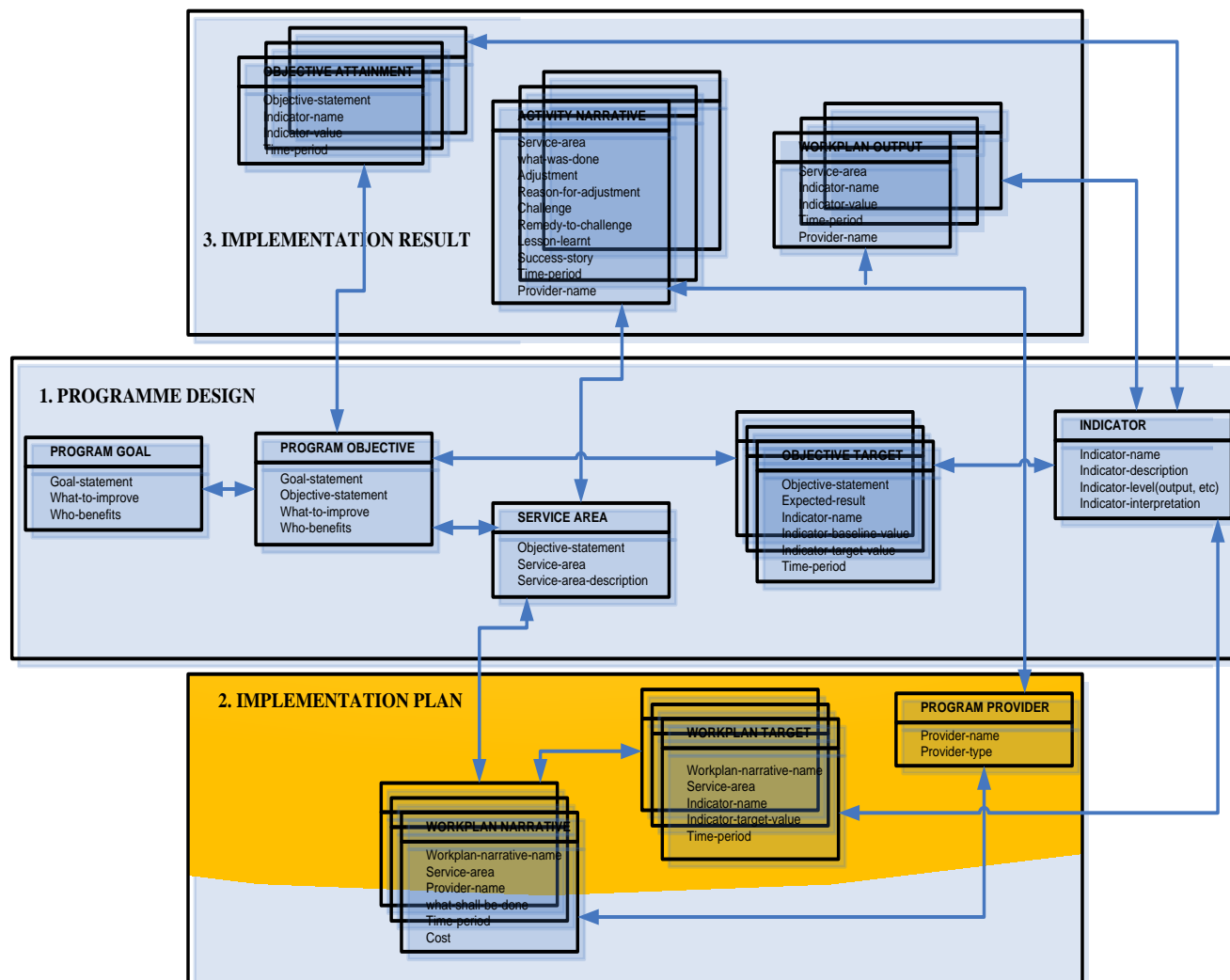


6.6 The Reflector Conceptual Model

This concluding section synthesizes the distinctions discussed in the previous sections. The outcome of the synthesis is a conceptual model, which has been named **Reflector** – a name chosen to highlight the fact that it “mirrors” both the practice and theory of M&E. The model has been created based on a set of eleven distinctions derived from the discussions of the previous sections. The attributes and relationships among the eleven distinctions of the model are shown in the diagram below. As portrayed in the diagram; the eleven distinctions have been arranged into a framework comprising of three dimensions: (1) programme design; (2) implementation plan; and (3) implementation result.

The *Programme Design* dimension comprises of five distinctions namely; PROGRAMME GOAL; PROGRAMME OBJECTIVE; SERVICE AREA; OBJECTIVE TARGET AND INDICATOR. These distinctions are used to capture the essential information relating to a programme design. The premise is that each PROGRAMME GOAL is operationalized by one or more PROGRAMME OBJECTIVE that specified the intention of the programme in actionable terms. This has often culminated into a set of one or more SERVICE AREA defined for each PROGRAMME OBJECTIVE. The SERVICE AREA specifies the products and services that will result from the implementation of the various PROGRAMME OBJECTIVES. Each PROGRAMME OBJECTIVE is also linked to one or more OBJECTIVE TARGET that specifies the minimum level of achievement or results expected from the objective. This minimum level of achievement or results is represented by an attribute named expected-result. The minimum level of achievement or results may relate to an output; outcome; or impact. Each OBJECTIVE TARGET is specified in terms of INDICATOR, which provides a mean of measuring the results specified in the expected-result attribute.

Figure 6.6: The Reflector Conceptual Model



The *Implementation Plan* dimension comprises of three distinctions namely; WORKPLAN NARRATIVE; WORKPLAN TARGET and PROGRAMME PROVIDER. These distinctions capture the essential information on the actions that have been planned by each programme implementer. The premise is that after the design phase; a programme enters into the implementation phase. At implementation phase; several PROGRAMME PROVIDER are recruited to deliver the pre-specified programme services to beneficiaries. At specified intervals; each PROGRAMME PROVIDER prepares a WORKPLAN NARRATIVE that specifies the details of what it intends to do to deliver one or more of the services specified in the SERVICE AREA. In addition; the PROGRAMME PROVIDER also prepares the WORKPLAN TARGET that specifies the minimum level of achievement it expects to attain. Each WORKPLAN TARGET is specified in terms of INDICATOR, which details the dimensions on which achievement is measured.

The *Implementation Result* dimension comprises of three distinctions namely; OBJECTIVE ATTAINMENT; ACTIVITY NARRATIVE and WORKPLAN OUTPUT. These distinctions capture the essential information on the outcome of both routine and terminal monitoring and evaluation activities. The OBJECTIVE ATTAINMENT distinction documents the extent to which each programme objective has been achieved. This measurement is facilitated by the use INDICATOR that specifies the actual dimensions used for measuring the results. Observed results are sometimes compared with those that were specified in OBJECTIVE TARGET. This comparison has been made possible by having the two distinctions refer to the same set of INDICATORS. Likewise, the WORKPLAN OUTPUT distinction describes the extent to which the PROGRAMME PROVIDER's targets that were established within the WORKPLAN TARGET have been achieved. This measurement is facilitated by the use INDICATOR that specifies the actual dimensions used for measuring the achievement. Observed results are sometimes compared with those that were specified in WORKPLAN TARGET since both distinctions employ the same INDICATOR to measure progress. The ACTIVITY NARRATIVE captures qualitative descriptions of what unfolded during the implementation of the WORKPLAN NARRATIVE.

Chapter 7: Conclusions and recommendation

This thesis is grounded on a pragmatic observation that existing software for programme monitoring and evaluation are mainly designed to address the M&E information needs of distinct programme or project, and are less adaptable to the information needs of alternative programmes or projects. This situation has been exacerbated by the lack of studies that focus on arranging the various concepts associated with programme monitoring and evaluation activities into a coherent pattern of concepts – vis-à-vis conceptual model. In this thesis, it is argued that availing such literature could stimulate the development of generic off-the-shelf software for programme M&E. It is observed that disciplines such as project management, accounting and constructions have benefited from such conceptual models as evidenced by the existence of several off-the-shelf project management software packages; and financial accounting packages. Consequently, the empirical part of this thesis seeks to review the practice of monitoring and evaluation within three large scale programmes with a view to identify and characterize the key concepts used in each case study programme. This aspect of the study is guided by an analytical framework developed in the thesis. The analytical framework reveals a set of key concepts that characterize any programme monitoring and evaluation initiative. These concepts are searched for within the three case study programmes and characterized the result in terms of identified attributes and observed relationships. Through further analysis, a conceptual model for programme monitoring and evaluation information system: the “Reflector” model basing on the outcome of the empirical study is developed. In the sections below, the conceptual model developed in the thesis is discussed in the view of previous research in the area of programme monitoring and evaluation, particularly evaluation models.

Wasserman (2008) identifies and categorized the key distinctions that apply to the monitoring and evaluation of a human service programme¹⁹. She categorizes the distinctions according to the three systems that make up a human service programme: provider, target and supra systems. As reflected in the conceptual model developed and presented in the previous chapter, not all the key distinctions that Wasserman (2008) identified were identified in the three case study programmes. In effect, the Reflector model is developed from just a few of the distinctions: PROGRAMME ACTIVITY, GOAL, OBJECTIVE, OUTCOME, PROVIDER, and TARGETED BENEFICIARY. Furthermore, two of the distinctions; PROGRAMME OUTCOME and TARGETED BENEFICIARY are modeled as attributes of some other distinction. For instance, PROGRAMME OUTCOME is modeled as the *expected-result* attribute of the OBJECTIVE TARGET distinction. Likewise, the TARGETED BENEFICIARY distinction is modeled as *who-benefits* attribute of the PROGRAMME GOAL and PROGRAMME OBJECTIVE.

¹⁹ In this thesis, it is argued that these same set of distinctions are also applicable to any programme other than just human service programme.

The fact that only a subset of the distinctions listed in Wasserman (2008) can be identified from the three case study programme resonates with the views of evaluation theorists that every evaluation situation is distinct, and needs tailoring to suit the purpose; the evaluator's preference of approaches and the nature of the evaluator-stakeholder relationship (Rossi, et al, 2004; Weiss 2005). In view of the distinctions, attributes and relationships portrayed in the Reflector model; it is plausible to conclude that the three case-study programmes, and indeed the Reflector model are focused at producing evidence to serve the accountability requirements of programme stakeholders. Although Rossi and Freeman (1987:157f, cited in Vedung, 1997) lists several dimensions of accountability, the Reflector model seems aligned more to the delivery and impact dimensions. Delivery accountability seeks to answer questions such as *are proper amounts of outputs being delivered? Are the treatments delivered those the program[me] is supposed to be delivering?* In the Reflector model, these questions are addressed through the following distinctions: ACTIVITY NARRATIVE, WORKPLAN NARRATIVE, WORKPLAN OUTPUT, WORKPLAN TARGET, and INDICATOR. The WORKPLAN TARGET distinction captures the amount of output that the programme expects to achieve within a specified time-period. Likewise, the WORKPLAN OUTPUT distinction captures the actual amount of output that the programme has produced within a specified period of time. Each output is measured using a set of indicators defined in the INDICATOR distinction. Inference regarding whether the proper amount of output was produced or not is derived by assessing the outputs captured in the WORKPLAN TARGET against WORKPLAN OUTPUT distinctions. Likewise, inference regarding whether the treatments delivered are those the programme is supposed to deliver is realized by comparing the details of the WORKPLAN NARRATIVE with ACTIVITY NARRATIVE. A useful aspect of the Reflector model is the ability to categorize the various inferences according to the PROGRAMME PROVIDER responsible for its attainment. In regards to impact accountability, the main question sought is whether *the programme is producing the intended outcome?* In the Reflector model, this question is addressed by the following distinctions: INDICATOR, OBJECTIVE ATTAINMENT, and OBJECTIVE TARGET. The OBJECTIVE ATTAINMENT distinction captures the details of the outcomes produced by the programme.

The OBJECTIVE TARGET captures the details of the outcomes that the programme expects to achieve. Inference regarding whether the programme is producing the intended outcome is realized by assessing the outcomes captured in the OBJECTIVE ATTAINMENT against OBJECTIVE TARGET distinctions. Each outcome is measured using a set of indicators defined in the INDICATOR distinction. A disclaimer is needed here though: it is acknowledged that several aspects of delivery and impact accountability M&E may actually fall outside the scope of the Reflector model. For instance, the Reflector model is most suited to impact monitoring (showing evidence of the existence of outcome) than impact assessment (addressing causation in relation to observed outcome).

This thesis begins with the ambition to develop a conceptual model for a programme monitoring and evaluation information system that will stimulate development of generic software for monitoring and evaluating programmes. Through the review of literature on evaluation models; the field of evaluation is found to be besieged with disagreement and lack of consensus on the very basic issues such as nomenclature and definitions. Rather than base the development of the conceptual model on just the theory of evaluation, the practice of evaluation within three large scale programmes was used to compliment theory with practice. This complementary process of evolving the Reflector model enables a comparison of the conceptual model developed in the study with existing theories of evaluation practice. As discussed above, the conceptual model developed in this study addresses many of the accountability requirements of programme stakeholders. In addition, it's essential structure matches with many of the accountability-oriented studies discussed in the literature review chapter (chapter two). Through this study a model of evaluation that combines functionalities of several evaluation approaches is developed. It is expected that this work can stimulate interest in studies that result into several conceptual models similar to the one developed but that address other equally important facets of evaluation. The practical significance of this study is that the Reflector model can be translated into a software package that may be useful to simplify data management and reporting for organizations concerned with issues of programme accountability. Even in its current state, the REFLECTOR model specifies constructs and parameters that may guide organizations on the complete array of issues that need to be addressed in evaluation planning and reporting.

The model developed in this study is a valuable artifact that provides a comprehensive knowledge base for planning and executing accountability-oriented evaluation studies. The reflector model reduces the many theories and concepts of evaluation into a set of eleven core concepts that connects both the theories and the practice of evaluation. Focusing evaluation on just a core set of eleven concepts reduces the amount of effort required to collect, store, analyze and report on evaluation data. This is likely to reduce the burden of excessive data collection usually placed on programme implementers while fostering compliance to data collection routines. The application of this model is best done within a computerized information system. With just a core set of eleven concepts, programme agencies are able to develop computerized databases that automate the capturing, storage, and reporting of evaluation data on the eleven concepts of the Reflector model. A computerized database would provide a value-added service in terms of ensuring uniformity in planning, monitoring and evaluation in all programme sites.

REFERENCES

1. Abma, T. & Stake, R., E. (2001). Stake's Responsive Evaluation: Core ideas and Evolution. *New Directions for Evaluation*, 92:7-22.
2. Ahlemann, F. (2009). Towards a conceptual reference model for project management information systems. *Project Management*, 27, 19-30.
3. Ahlemann, F., Gastl, H. (2007). Process Model for an Empirically Grounded Reference Model Construction. In P. Fettke & P. Loos (Ed), *Reference Modeling for Business Systems Analysis* (pp. 77-97). Hershey: Idea Group.
4. Aid-IT Solutions. (2007). Aid-IT Solutions – M&E Professionals. [World Wide Web] Available: (<http://aid-it.com.au>) [2009, July 26].
5. Alexa, A., & Zuell, C. (1999). *A review of software for text analysis*. Mannheim, Germany: ZUMA.
6. Anderson, V., & Johnson, L. (1997). *Systems thinking basics: From concepts to causal loops*. Waltham, MA: Pegasus Communications, Inc.
7. Babbie, E. R., & Mouton, J. (2001). *The practice of social research*. Cape Town: Oxford University Press Southern Africa.
8. Bawden, R. (2006). A Systemic Evaluation of an Agricultural Development: A Focus on the Worldview Challenge. In B. Williams and I. Imam (Ed), *Systems Concepts in Evaluation: An Expert Anthology* (pp. 35–47). Point Reyes: EdgePress of Inverness.
9. Branda, N. (2008). Response to paper “Systems Thinking” by D. Cabrera et al.: Conceptualizing systems thinking in evaluation. *Evaluation and Program Planning*, 31: 311–334.
10. Bulgarelli, A., & Gori, E. (2004). Information Systems for the Evaluation of the Effectiveness and Efficiency of Vocational Training Programmes. *Evaluation*, 10(2):217-235.
11. Cabrera, D., Colosi, L., & Lobdell, C. (2008). Systems thinking. *Evaluation and Program Planning*, 31(3): 299–310.
12. CDC. (2008). PEMS Data Variables and Values. [World Wide Web] Available: (http://www.ct.gov/dph/lib/dph/aids_and_chronic/pems/pems_data_dictionary.pdf) [2009, April 24].

13. Chalmers, N.A. (2005). A Review of the Ugandan Local Government Information and Communication System – LoGICS: Consultancy report/ prepared for DFID by Neil Chalmers.
14. CHAMP. (2005). *Bush Administration Imposes Sweeping Data Collection on U.S. HIV Prevention Programs*. [World Wide Web] Available: (<http://www.champnetwork.org/media/PEMSDec20.pdf>) [2009, April 26].
15. Cook, T.D., & Shadish, W.R. (1986). Program Evaluation: The Worldly Science. *Annual Review of Psychology*, 37, 193-232.
16. Crawford, P. (2004). Aiding Aid: A monitoring and evaluation framework to enhance international aid effectiveness. Unpublished doctoral thesis, UTS, Sydney.
17. Crawford, P., & Bryce, P. E. (2003). Project monitoring and evaluation: a method for enhancing the efficiency and effectiveness of aid project implementation. *International Journal of Project Management*, 21:363–373
18. Crawford, S. (2002). Evaluation of Web survey data collection systems. *Field Methods*, 14 (3), 307–21.
19. Crofts, N., Doerr, M. & Gill, T. (2003). The CIDOC Conceptual Reference Model: A standard for communicating cultural contents. *Cultivate Interactive*, issue 9. [Electronic] Available: <http://www.cultivate-int.org/issue9/chios/index.htm> [2009, August 7].
20. Dieste, O., Juristo, N., Moreno, A.M, Pazos, J., & Sierra, A. (2000). Conceptual modelling in software engineering and knowledge engineering: concepts, techniques and trends. *Handbook of Software Engineering and Knowledge Engineering*. World Scientific
21. DoView (2009). The DoView User Manual. [World Wide Web] Available: (<http://www.doview.com/files/helppdf/doviewhelpmanual2-0.pdf>) [2009, August 7].
22. Fettke, P., & Loos, P. (2003). Classification of reference models: a methodology and its application. *Information Systems and e-Business Management*, 1:35-53.
23. Fettke, P., & Loos, P. (2007). Perspectives on Reference Modeling. In P. Fettke & P. Loos (Ed), *Reference Modeling for Business Systems Analysis* (pp. 1-20). Hershey: Idea Group.
24. Hansen, H. F., (2005). Choosing Evaluation Models. *Evaluation*, 11(4), 447–462.
25. Imam, I., and Williams, B. (2007). *Systems Concepts in Evaluation: An Expert Anthology*. Point Reyes: EdgePress of Inverness.

26. Imam, I., LaGoy, A., & Williams, B (2006). Introduction. In B. Williams and I. Imam (Ed), *Systems Concepts in Evaluation: An Expert Anthology* (pp. 3–10). Point Reyes: EdgePress of Inverness.
27. InstantSurvey. (2009). Online Survey Software tool for Easy Online Surveys: InstantSurvey Online Survey Software. [World Wide Web] Available: (<http://www.instantsurvey.com/>) [2009, March 25].
28. Jackson, M.C., (2003). *Systems Thinking: Creative Holism for Managers*. Point West Sussex: John Wiley & Sons.
29. Juristo, N., Moreno, A.M. (2000). Introductory paper: Reflections on Conceptual Modelling. *Data & Knowledge Engineering*, 33, 103-117.
30. LaPelle, N.R. (2004). Simplifying qualitative data analysis using general purpose software tools. *Field Methods*, 16(1), 85–108.
31. Leslie, D.R, Holosko, M.J., & Dunlop, J.M. (2006). Using Information Technology in Planning program Evaluation. *Journal of Evidence-Based Social Work*, 3(3/4), 73–90
32. Mark, M. (2005). Evaluation Theory or What Are Evaluation Methods for? *Evaluation Exchange*, 11(2), 2-3.
33. Mark, M. M., Henry, G. T., & Julnes, G. (2000). *Evaluation: An integrated framework for understanding, guiding, and improving policies and programs*. The Jossey-Bass nonprofit and public management series. San Francisco: Jossey-Bass.
34. MEEPP Project. (2008). *User's Guide: The MEEPP Web-based Reporting System*. [World Wide Web] Available: (<http://www.meepp.or.ug>) [2009, May 3].
35. Midgley, G. (2006). Systems Thinking for Evaluation. In B. Williams and I. Imam (Ed), *Systems Concepts in Evaluation: An Expert Anthology* (pp. 11–34). Point Reyes: EdgePress of Inverness.
36. Moon, T., & Smith, S. (n.d). MEEPP: Reporting on Results in Uganda. [World Wide Web] Available: (<http://www.globalhealth.org/reports/report.php3?id=243>) [2009, June 17].
37. Mouton, J. (2008). *How to succeed in your Master's and Doctoral Studies: A South African Guide and Resource Book*. Pretoria: Van Schaik Publishers
38. Niehaves, B., Ribbert, M., Dreiling, A., & Holten, R. (2004). Conceptual Modeling – An Epistemological Foundation. in AMCIS: *Proceedings of the Tenth Americas Conference on Information Systems* (pp. 4232-4242). Atlanta: Association for Information Systems.

39. PEPFAR. (2008). The President's Emergency Plan for AIDS Relief: indicators, reporting requirements, and guidelines. [World Wide Web] Available: (<http://www.pepfar.gov/documents/organization/113759.pdf>) [2009, August 7].
40. Rossi, P.H., Lipsey, M.W., & Freeman, H.E. (2004). *Evaluation: A Systematic Approach*. California: Sage.
41. Rossi, P.H., Lipsey, M.W., & Freeman, H.E. (2004). *Evaluation: A Systematic Approach*. Thousand Oaks, Calif. [u.a.]: Sage Publ.
42. Scriven, M., (2003). Evaluation Theory and Metatheory. In D. L. Stufflebeam, & T. Kellaghan, (Eds.), *the international handbook of educational evaluation* (Chapter 1). Boston: Kluwer Academic Publishers.
43. Smith, E., E. (1995). Concepts and categorization. In: Kosslyn, S.M, Osherson D, (Eds.), *an invitation to cognitive science, Vol. 2*: Cambridge (MA): MIT Press.
44. SPSS Inc. (2009). SPSS, Data Mining, Statistical Analysis Software, Predictive Analytics. [World Wide Web] Available: (<http://www.spssmr.com/>) [2009, March 20].
45. Stiemerling, O., Kahler, H., & Wulf, V. (1997). How to make software softer: Designing tailorable applications. In: *Proceedings of the ACM Symposium on Designing Interactive Systems* (pp. 365 – 376). Amsterdam: ACM-Press.
46. Stufflebeam, D. L. (2003). The CIPP model for evaluation. In D. L. Stufflebeam, & T. Kellaghan, (Eds.), *the international handbook of educational evaluation* (Chapter 2). Boston: Kluwer Academic Publishers.
47. Stufflebeam, D.L. (2001). *Evaluation models*. San Francisco: Jossey-Bass.
48. SurveyMonkey.Com. (2009). SurveyMonkey.Com – Powerful tool for creating web surveys. Online survey software made easy! [World Wide Web] Available: (<http://www.surveymonkey.com/>) [2009, March 22].
49. SysCorp International. (2006). Analytical report on the assessment of the sector information systems & assessment of the implication of interfaces with other existing sector systems. Consultancy report prepared for MOLG by SysCorp International.
50. Tesch, D., Sobol, M.G, Klein, G., & Jiang, J.J. (2009). User and developer common knowledge: Effect on the success of information system development projects. *International Journal of Project Management*, 27(7), 657-664.
51. Thomas, C. (2005). *Overview of HIV/AIDS Prevention Program Evaluation and Monitoring System*. Unpublished PowerPoint Slide. [World Wide Web] Available:

- (http://www.cdc.gov/hiv/topics/perinatal/resources/meetings/2005/pdf/Thomas_Craig_PE_MS.pdf) [2009, April 20].
52. Thomas, C.W., Smith, B.D., & Wright-De Agüero, L. (2006). The Program Evaluation and Monitoring System: A Key Source of Data for Monitoring Evidence-Based HIV Prevention Program Processes and Outcomes. *AIDS Education and Prevention*, 18(suppl. A), 74-80.
 53. UNAIDS. (2003). Country Response Information System: Overview of the system and its plan of establishment. [World Wide Web] Available: (http://data.unaids.org/Publications/IRC-pub02/jc885-cris_overview_en.pdf) [2009, May 9].
 54. UNAIDS. (n.d). Welcome to CRIS – Country Response Information System. [World Wide Web] Available: (<http://www.pepfar.gov/documents/organization/113759.pdf>) [2009, August 7].
 55. Vedung, E. (1997). *Public Policy and Program Evaluation*. London: Transaction Publishers.
 56. Vovici Corporation. (2009). Survey Software for Enterprise feedback Management & Online Community Software. [World Wide Web] Available: (<http://www.vovici.com/>) [2009, March 29].
 57. Wand, Y., Monarchi, D. E., Parsons, J., & Woo, C. C. (1995). Theoretical foundations for conceptual modelling in information systems development. *Decision Support Systems*, 15, 285-304
 58. Wasserman, D., L. (2008). Response to paper “Systems Thinking” by D. Cabrera et al.: Next steps, a human service program system exemplar. *Evaluation and Program Planning*, 31: 327–329.
 59. Wasserman, D., L. (2010). Using a systems orientation and foundational theory to enhance theory-driven human service program evaluations. *Evaluation and Program Planning*, 33(2010): 327–329.
 60. Weber, R. (2003). Conceptual Modelling and Ontology: Possibilities and pitfalls. *Database Management*, 14(3): 1-20.
 61. Whitten, J.L., Bentley, L.D., & Dittman, K.C. (2004). *Systems analysis and design methods*. New York: McGraw Hill.
 62. Williams, B., & Iman, I. (2006). *Systems concepts in evaluation: An expert anthology*. Point Reyes, CA: EdgePress.

63. Wright, K. (2005). Researching Internet-based populations: Advantages and disadvantages of online survey research, online questionnaire authoring software packages, and web survey services. *Journal of Computer-Mediated Communication*, 10(3) [Electronic] Available: <http://jcmc.indiana.edu/vol10/issue3/wright.html> [2009, April 25].
64. Wyssusek, B. (2006). On Ontological Foundations of Conceptual Modelling. *Scandinavian Journal of Information System*, 18(1), 63-80.
65. Zoomerang. (2009). Zoomerang – Easiest Way to Ask, Fastest Way to Know. [World Wide Web] Available: (<http://www.zoomerang.com/>) [2009, March 18].